Design and Synthesis of Peptides Derived from Nucleoside β-Amino Acids

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

By

Michelle Louise Turtle

2012

Abstract

Naturally occurring macromolecules such as proteins and DNA adopt very specific conformations and three dimensional arrangements which allow them to perform their sophisticated functions in nature. Unnatural oligomers can also adopt well defined conformations and these have been termed 'foldamers'. This project is concerned with the design and synthesis of β-peptide foldamers which are assembled from nucleoside derived β-amino acids. It was hoped that the combination of the inherent helix folding properties of peptides and the associated characteristics of nucleosides would allow us to create novel foldamers with specific recognition properties. The structure of the β-amino acids involves conversion of the 3′-hydroxyl group of the nucleoside sugar into an amino group (with retention of configuration) and oxidation of the 5′-hydroxymethylene group to a carboxylic acid.

Beginning with the natural nucleosides thymidine and 2′-deoxyadenosine, syntheses of 4 different nucleoside β-amino acid monomers was accomplished. The syntheses were based around the use of 3′-azido-2′-deoxynucleosides as key synthetic intermediates and also required optimisation of protecting groups to prevent side reactions associated with cleavage of the glycosidic bond. In general, the preferred protecting group strategy involved the 3′-amino group being either protected with a fluorenylmethoxycarbonyl (Fmoc) group or being masked as an azide. The benzhydryl group also proved useful in the protection of the carboxylic acid function for couplings in solution.

The synthesis of the β-peptides was investigated by both solid-phase peptide synthesis and standard couplings in solution. A range of peptide dimers, trimers and a tetramer were prepared, but the isolated yields of the final products were low due to difficulties with purification.

Contents

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Professor Rick Cosstick. Without your guidance, encouragement and unrivalled enthusiasm for the project I don't think I would have got to where I am today and for that I am eternally grateful.

I would also like to thank all the past and present members of the Cosstick/Carnell lab group for making the lab such a friendly place to work! In particular I would like to thank Dr. James Gaynor and Dr. Ralph Kirk who are two of the most hardworking people I know! Thank you for being great friends and tolerating my stupid questions and complaining at tea breaks!

A big thank you to Dr Julie Fisher at the University of Leeds for sharing her expertise in NMR and also to all the technical staff at the University of Liverpool. In particular I would like to thank Alan Mills and Moya McCarron who both work very hard on the university mass spectrometry service.

I would like to acknowledge the Engineering and Physical Sciences Council for providing the generous funding for this project.

Finally I would like to thank my family and friends for all their support and encouragement over the past 4 years. Most of all I would like to say how grateful I am to my fiancé Robin, I don't know where I would be without you. I cannot thank you enough for the emotional support you have given me over the past 4 years and I wish to dedicate this thesis to you.

Abbreviations

A	Adenine
AA	Amino acid
Ac_2O	Acetic anhydride
ACHC	2-Aminocyclohexanecarboxylic acid
ACPC	2-Aminocyclopentanecarboxylic acid
Alloc	Allyloxycarbonyl
AZT	3'-Azido-3'-deoxythymidine
B	Nucleobase
BAIB	Bis-acetoxyiodobenzene
BOC	tert-Butoxycarbonyl
BOP	Benzotriazol-1-yloxytris(dimethyl-amino)-phosphonium hexafluorophosphate
Bz	Benzoyl
C	Cytidine
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	Dicyclohexyl carbodiimide
DCM	Dichloromethane
DIAD	Di*iso*propylazdicarboxylate
DIC	Di*iso*propylcarbodiimide
DIPA	Di*iso*propylamine
DIPEA	Di*iso*propylethylamine
DMAc	*N,N*-Dimethylacetamide
DMAP	Dimethylaminopyridine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethylsulfoxide
DMT	4,4'-Dimethoxytrityl
DNA	Deoxyribonucleic acid
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ES	Electrospray
Et	Ethyl
Et_2O	Diethyl ether
EtOAc	Ethyl acetate
ETT	5-(Ethylthio)-1H-tetrazole
Fmoc	9-Fluorenylmethoxycarbonyl
G	Guanine

HATU	*O*-(7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	*O*-(Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HIV	Human immunodefficiency virus
HMPA	Hexamethylphosphoramide
HMPB-BHA	4-Hydroxymethyl-3-methoxyphenoxybutyric acid benzhydrylamine
HMPB-MBHA	4-Hydroxymethyl-3-methoxyphenoxybutyric acid 4-methylbenzhydrylamine
HOAt	7-Aza-1-*N*-hydroxy benzotriazole
HOBt	1-*N*-Hydroxy benzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IR	Infrared
m-	*Meta*
MALDI-TOF	Matrix assisted laser desorption ionisation-time of light
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
MHz	Megahertz
NaOMe	Sodium methoxide
NaOH	Sodium hydroxide
NMR	Nuclear magnetic resonance
o-	*Ortho*
OAt	7-Azabenzotriazolyloxy
OBt	Benzotriazolyloxy
OMs	Methanesulfonyl
p-	*Para*
PAC	Phenoxyacetyl
PG	Protecting group
Ph	Phenyl
PMBz	*para*-Methoxybenzoyl
PNBz	*para*-Nitrobenzoyl
PNA	Peptide nucleic acid
ppm	Parts per million

PyAOP	[(7-Azabenzotriazol-1-yl)oxy]*tris*-(pyrrolidino)-phosphonium hexafluorophosphate
PyBOP	Benzotriazol-1-yloxytri(pyrrolidino)-phosphonium hexafluorophosphate
Pyr	Pyridine
RNA	Ribonucleic acid
SPPS	Solid phase peptide synthesis
T	Thymine
TBAF	Tetrabutylammonium fluoride
TBDMS	*tert*-Butyldimethylsilane
TBDPS	*tert*-Butyldiphenylsilane
TEAB	Tetraethylammonium bicarbonate
TEMPO	2,2,6,6-Tetramethylpiperidinyloxy
TFA	Trifluoroacetic acid
TfO	Trifluoromethanesulfonic
THF	Tetrahydrofuran
TIPS	Tetra*iso*propyldisiloxane
tlc	Thin layer chromatography
TMS	Tetramethylsilane
TNBS	2,4,6-Trinitrobenzenesulfonic acid
U	Uracil
UV	Ultraviolet

Chapter 1: Introduction

1.1 Introduction

Within the natural world, a large proportion of complex processes are carried out by macromolecules. Each of these molecules adopts a highly ordered structure which is specifically adapted to a unique role. These macrostructures are extremely kinetically and thermodynamically stable and this stability can be attributed to the presence of a high number of H-bond donors and acceptors in most cases. Some of the most important examples of these highly ordered structures are peptides and deoxyribonucleic acid (DNA); both these molecules exhibit high levels of hydrogen bonding including inter- and intra- molecular bonding which attributes to and allows them to keep their distinct structures. Over the past 60 years, the interest in these types of molecular structures and their mimics has proved to be a major and important area of research.

1.2 DNA Structure

1.2.1 Primary structure

Although DNA was first isolated in the late 1800's,[1] its true biological function was only demonstrated by Avery and co-workers in 1944.[2] Early studies of the molecule had given a basic understanding of the chemical content of DNA.[3,4] However very little was known about its higher order structure and how this contributed to its function until the publication by Watson and Crick of its double helix structure in 1953.[5] This realisation that structure, as well as the chemical content of a molecule, was important for biological activity was the starting point for an extremely interesting area of research.

In order to explore the higher order structures of DNA, the primary structure must first be examined. The features of the individual nucleotide units are ultimately responsible for the complex helical structure formed by DNA and each unit consists of a furanose sugar, a heterocyclic base and a phosphate ester.

The bases are either monocyclic pyrimidines (cytosine and thymine, uracil in the case of ribonucleic acid) or bicyclic purines (guanine and adenine) and these hydrogen bond with each other in a highly specific manner. Adenine is always paired with thymine (or uracil in RNA) through 2 hydrogen bonds and guanine with cytosine through 3 hydrogen bonds, making this the stronger of the 2 base pairs (Figure 1.1). Watson-Crick base pairing is the dominating pattern in double stranded

nucleic acid structures. However, other hydrogen bonding arrangements such as Hoogsteen pairing are often found in multi-stranded structures (Figure 1.1).[6,7]

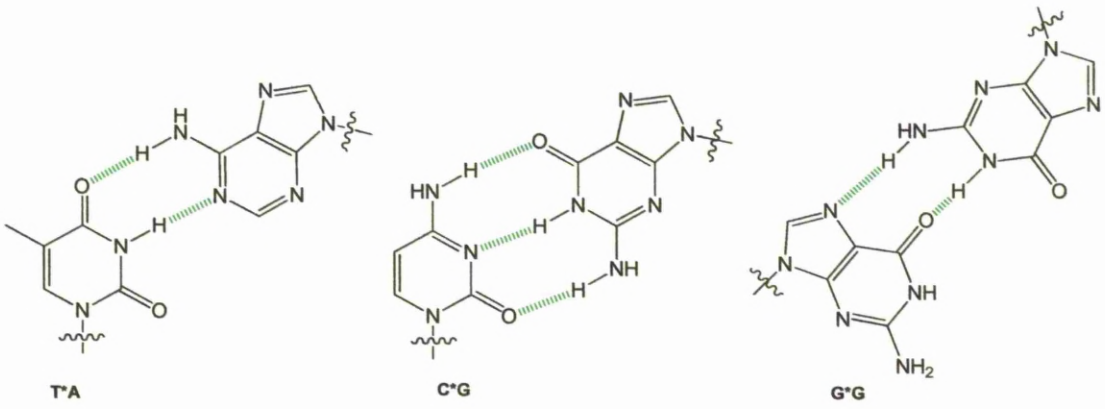


Figure 1.1: Watson-Crick base pairing for T-A (left), C-G (middle) and Hoogsteen base pairing for G-G as seen in G-quadruplexes (right).

The heterocyclic base is attached to the furanose sugar at the anomeric centre (C1′) and exclusively adopts the β-configuration. The pyrimidine bases are connected to the sugar by the *N*-1 atom and the purines by *N*-9. The bond formed between the sugar and base is termed the glycosidic bond and the molecule is collectively known as a nucleoside. When esterified with a phosphate ester at the 5′-hydroxyl group, these compounds are then termed nucleotides (Figure 1.2). These monomers are then joined together in a manner where a phosphodiester bond is formed between the 5′-position of one and the 3′-position of its neighbour.

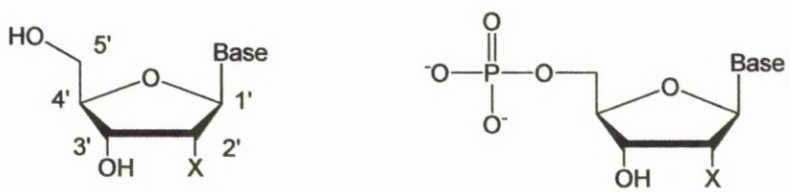


Figure 1.2: A Nucleoside (left) a Nucleotide (right) In DNA X=H, in RNA X=OH.

The major structure defining features of the nucleosides are the pucker adopted by the furanose sugar ring and the position of the base with respect to the rest of the sugar. Studies into nucleoside structure based on pseudorotations have identified two interconverting sugar conformations which are termed the C3′-*endo* conformation and the C2′-*endo* conformation (Figure 1.3). The *endo* terminology refers to the displacement of the C2′ or C3′ atom out of the C1′-O4-C4′ plane of the furanose ring, with the *endo* face being the side on which the heterocyclic base

resides. The other side is termed the *exo* face. The terms 'North' (N) and 'South' (S) are also used to describe the C3'-*endo* and C2'-*endo* conformations, respectively and are based upon their geographical positions on the pseudorotational cycle.[8] Conveniently, the 'N' and 'S' notation also represents the approximate shape of the C4'-C3'-C2'-C1' bonds in each conformation.

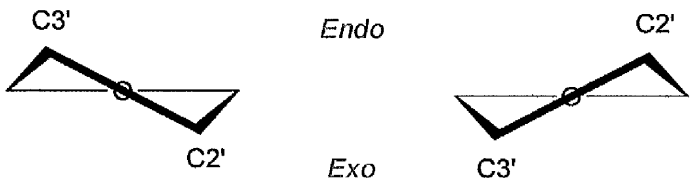


Figure 1.3: Simplified furanose sugar puckers C3'-*Endo* "North" (left) and C2'-*Endo* "South" (right).

The preferred sugar conformation of a nucleoside is determined by a number of factors including the orientation of the base and electronic effects in the ring system itself. In DNA, the south-type sugar pucker is generally preferred.[9] A major contributing factor in the adoption of the south conformation is a favourable gauche interaction (60° dihedral angle when looking along the C3'-C4' bond) between the ribose ring oxygen (O4') and the oxygen presiding at C3' (O3').[10]

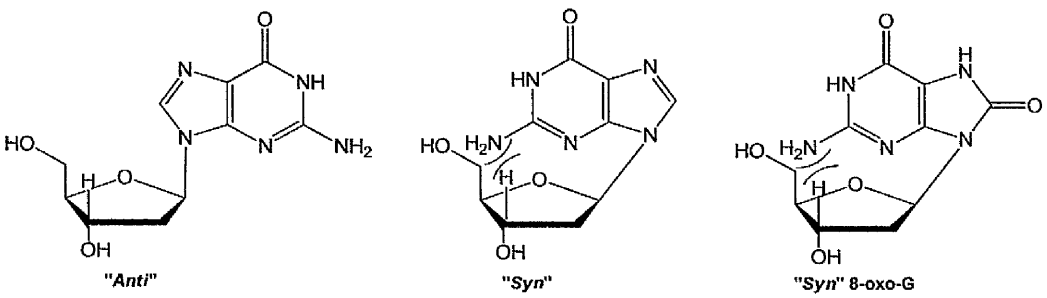


Figure 1.4: Nucleoside conformations of guanosine: sterically favourable "*anti*" (left), sterically disfavoured "*syn*" (middle) and favoured "*syn*" conformation for 8-oxo-G (right).

The bases in nucleosides can be orientated in two directions with respect to the sugar and are termed '*anti*' and '*syn*' (Figure 1.4). The '*anti*' conformation is usually preferred and describes the situation when the major bulk of the base is orientated away from the sugar ring; this reduces any unfavourable steric interactions. This conformation is adopted exclusively by the pyrimidines choosing to have the smaller H6 above the plane of the ring and avoiding any steric clash between the bulkier O2 and H3'. Adenine also prefers the '*anti*' conformation. However, it has been shown

that, in certain cases, guanine can prefer to adopt the '*syn*' conformation. It has been shown that substitutions at the 8-position of the guanine base, such as the formation of 8-oxo-G in oxidatively damaged DNA,[11] can result in a favouring of the '*syn*' conformation. The introduction of halogens at C8 to give 8-bromo-G,[12] 8-chloro-G and 8-iodo-G[13] also causes the nucleoside to favour the '*syn*' conformation. This characteristic of guanine can result in the formation of unusual structures in G-rich sequences, such as in G-quadruplexes which will be explained in more detail later in this chapter.

1.2.2 Secondary structure

The original structure of DNA, proposed by Watson and Crick, was a right-handed double helix comprised of two single strands of DNA running anti-parallel and held together by the formation of hydrogen bonding between the bases on opposite strands. As well as base pairing, an additional stabilising effect is π-π stacking between the heterocyclic bases in the helix (Figure 1.5). [14]

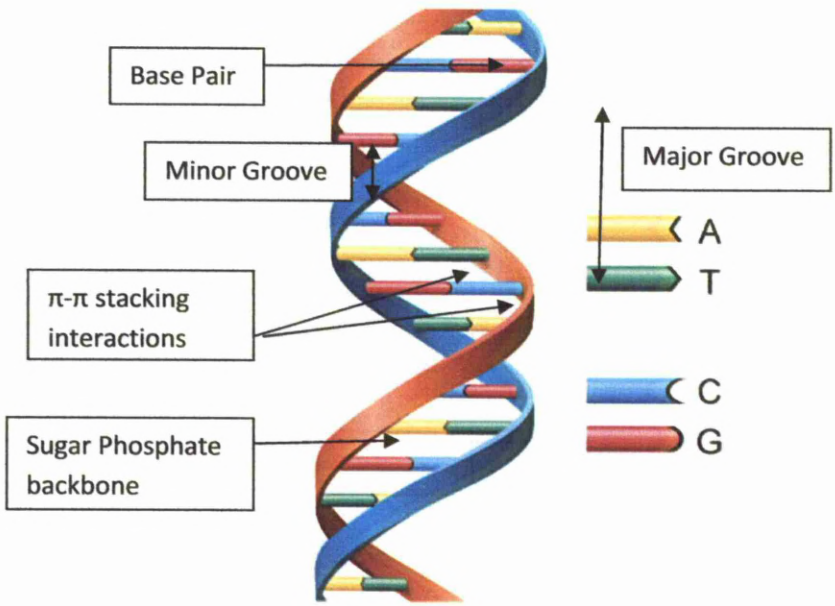


Figure 1.5: A double DNA helix (adapted from reference 15).[15]

In the 1970's, the development of automated nucleotide synthesis allowed for a huge leap in oligonucleotide structural studies. It became possible to obtain oligonucleotides of defined sequence and high enough purity for preparation of crystals allowing single crystal diffraction studies to be performed.[16,17]

At this time, it was already known that DNA exhibited a great deal of structural diversity. However, this was the first time where these structures, termed A-, B- and Z- forms could be studied in detail (Figure 1.6). Although different, these helical forms all possess the same structural features; a major groove and a minor groove which run down the helical axis and are opposite to each other.[18]

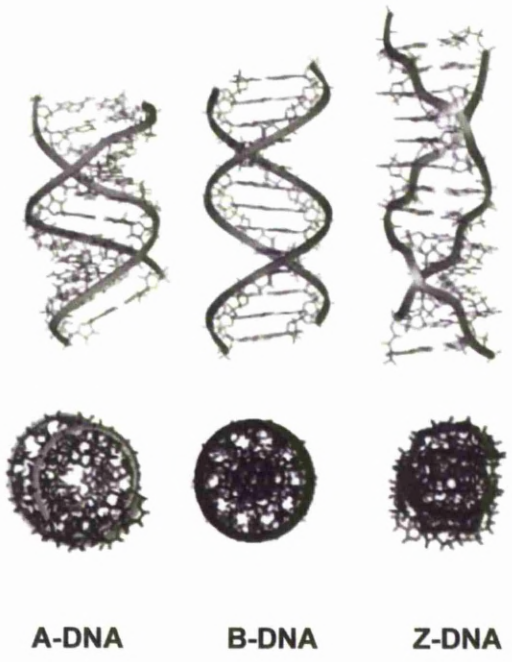


Figure 1.6: A-, B- and Z-DNA forms showing a side on view of the helix (top) and a view down the helical axis (bottom). The sugar phosphate backbone is shown as a ribbon (taken from reference 18).[18]

A-DNA was first observed in dehydrated DNA samples[19, 20] and it is a right-handed helix consisting entirely of C3'-*endo* nucleoside units with 11 base pairs per helical turn. The major groove of this DNA type is comparatively slim, at 2.7Å, but it is extremely deep (13.5Å). This slim wide major groove produces a wide shallow minor groove, having a width of 11.0Å but a depth of only 2.8Å. In A-DNA, the stacked base pairs are tilted slightly sideways in order to maintain a normal van der Waals' separation this results in a helix which is 26Å across.[14] A-DNA can be found in RNA duplexes[21] and DNA-enzyme complexes.[22]

B-DNA, also a right-handed helix, is the predominating form under standard physiological conditions and, unlike A-DNA, C2'-*endo* nucleoside units form the majority of the structure. In contrast to A-DNA, B-DNA has a wide major groove and slim minor groove. The major groove has a width and depth of 11.7Å and 8.8Å, respectively, whereas the minor groove has a width and depth of 5.7Å and 7.5Å,

respectively. This structure has 10 base pairs per helical turn and a diameter of 20Å.[14]

Z-DNA, a left-handed helix, is the most uncommonly seen form of DNA and it adopts an unusual structure. The true function of Z-DNA is not yet fully understood and it is regarded as more of a transient feature in a cellular environment than a permanent one.[23] It has been postulated that Z-DNA plays a role in the regulation of gene transcription[24] and a number of Z-DNA binding proteins have also been identified, including the EL3 protein which is responsible for the pathogenicity of poxviruses.[25]

The A-, B- and Z-forms of DNA, mentioned above, are the most intently studied and biologically relevant structures. However, the diversity of DNA extends further. A whole range of other DNA forms have been observed, although they are usually formed under very specific conditions or through specific chemical modifications. For example, L-DNA is formed when the natural deoxyribose sugar is replaced with L-deoxyribose. This imparts a resistance to nucleases into the DNA structure[26] and is sometimes termed 'mirror-DNA' as it is the mirror image of natural DNA. H-DNA is an intramolecular triple helical structure that forms under low pH conditions. H-DNA is thought to play a role in gene expression and only occurs in poly(purine): poly(pyrimidine) rich strands.[27] The range of structures found in DNA is so diverse that there are only 5 letters remaining to describe new forms that may be found in the future.[28]

1.2.3 Higher order structures

As methods for examining the structure of nucleic acids improved a number of multiple stranded structures began to emerge. Interest has grown in these higher order structures as their biological function and relevance becomes more apparent. Of particular interest are G-quadruplexes and intercalated motifs (i-motifs).

G-quadruplexes (sometimes termed tetraplexes, Figure 1.7) are found in G-rich strands of DNA and are square-like conformations of four guanine bases, termed tetrads, interacting by Hoogsteen hydrogen bonding in the presence of a monovalent metal ion.[29,30] These tetrads can be formed by 4 or 2 interacting strands and can even be formed by intramolecular interactions within the same strand; the latter is the most commonly found topology in nature.[31] G-quadruplexes of particular

interest are those which form at the end of eukaryotic chromosomes termed telomeres. In humans, the telomeric sequence is single stranded, 100-200 nucleotides in length and is a TTAGGG repeating sequence.[32] Human telomeres are maintained by the enzyme telomerase and this has been implicated in up to 85% of cancers. However, formation of G-quadruplexes has been shown to inhibit the activity of this enzyme.[33] More recently it has been observed that G-quadruplexes also play a role in the control and regulation of gene transcription. Studies have shown that subjecting genetic sequences that contain G-quadruplexes to small molecules or proteins that specifically target the tetrads can disturb genetic transcription.[34] This makes G-quadruplexes a very attractive target for both biological and chemical research.

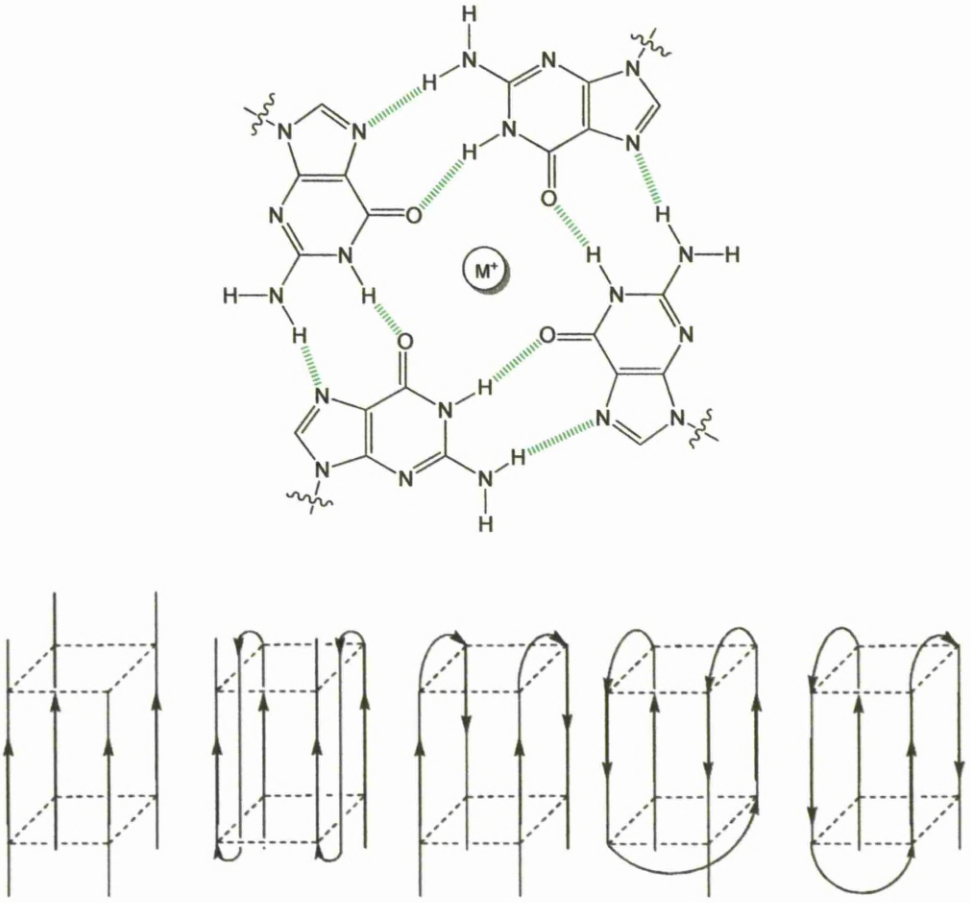


Figure 1.7: A G-tetrad (above) and folding topologies of G-quadruplexes (below) (adapted from reference 31).[31]

Interestingly, the C-rich strands complementary to the G-rich strands that form quadruplexes, can form their own complex higher order structure termed the i-motif.[35] Formation of the i-motif is pH dependent and consists of two parallel strands

which are held together by hemi-protonated C-C$^+$ base pairs.[36] Two duplexes can then interact in an anti-parallel fashion to form a complex quadruplex structure; this is the only known structure involving systematic base intercalation (Figure 1.8). As with the G-quadruplex the four interacting strands can be sourced from 4,[35] 2[37] or a single strand of DNA.[38] The biological function of the i-motif is not yet fully understood although recent research has shown that it too may have an important role in gene expression. A study of the oncogene *bcl*-2 found that deletion or mutation of a G-C rich region of the gene upstream from its P1 promoter site, which is known to play a large role in its transcription, caused an increase in transcription. This indicates that this G-C rich region and its ability to form higher order DNA structures must have a role in the regulation of gene expression.[39] Further research into this complex structure and its role in biology is of great interest in the scientific community and as techniques for the study of genetics become more advanced the information on the importance of higher order structures is being published.

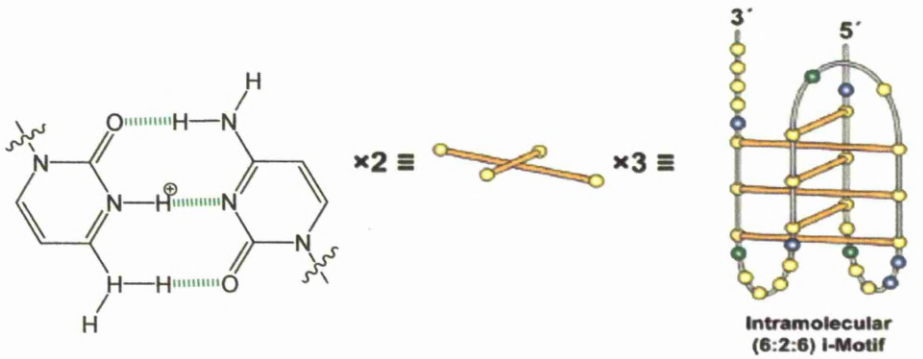


Figure 1.8: Hemi-protonated cytosine bases found in the i-motif (left) and a representation of a single stranded i-motif showing the systematic base intercalation (right), (adapted from reference 39).[39]

Higher order DNA structures are also being utilised in the area of nanotechnology where they have found a number of applications ranging from nanostructured materials to nanomachinery.[40, 41] G-Quadruplex motifs have been used in the formation of ion selective membrane channels[42] and DNA logic switches modulated by the presence of metal ions.[43] Both the G-quadruplex and i-motif have been employed in the formation of microfilms and microcapsules which, with further investigation, may have a use in selective binding of therapeutic proteins.[44] The i-motif alone has found applications in many areas, an example of which is the formation of nanoswitches such as those used in molecular machines known as DNA tweezers which are activated by a change in pH.[45] The research into DNA and

its structure is vast and diverse and continues to be of great interest in all areas of science and technology.

1.3 Peptide nucleic acid

1.3.1 History and primary structure

Peptide nucleic acid (PNA) is a DNA mimic in which the deoxyribose phosphate backbone has been replaced with a pseudo peptide backbone. In the original design of PNA (aegPNA), this backbone is made up of *N*-(2-aminoethyl)glycine units where the nucleobases are attached to the glycine nitrogen by a carbonyl methylene linker (Figure 1.9).

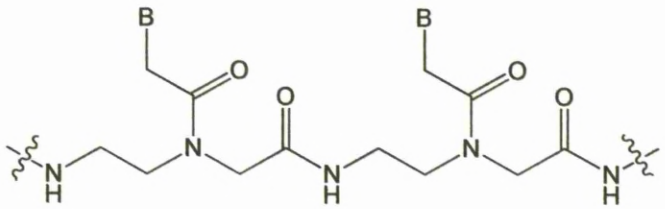


Figure 1.9: The original aegPNA backbone, B=nucleobase.

Since the introduction of aegPNA 20 years ago by Nielsen *et al.*,[46] there has been a vast amount of time and effort dedicated to this area in many fields of science from pure chemistry[47] and biology[48] to genetics[49] and nanotechnology.[50] PNA was originally designed as a DNA duplex binding ligand which would mimic a triplex-forming oligonucleotide and bind in the major groove of a DNA double helix through Hoogsteen type base pairing. It was designed based on the bonding pattern seen in oligonucleotides known as the "6 + 3 rule", that is, 6 bonds between neighbouring nucleotides and 3 bonds separating the backbone and heterocyclic base. It was found however that the PNA backbone was a much better substitute for a normal oligonucleotide backbone than expected.[46] The publication of this realisation sparked a frenzy of interest in this type of oligomer, particularly within the fields of antisense and antigene drug discovery.[51]

From a chemical point of view, PNA has a variety of interesting properties. PNA is acid stable and will not suffer from loss of the base even under the most harshly acidic conditions;[52] cleavage of the glycosidic bond is of major concern in DNA and its other mimics. PNA is stable to mildly basic conditions but will undergo acyl transfer reactions under strongly basic conditions.[53] For this reason, some of the preferred methods of PNA synthesis are the Boc- or Fmoc- solid phase strategies,

removal of Boc protection requires acidic conditions and Fmoc removal requires only extremely mild basic conditions.[54]

From a biological point of view PNA is an attractive target because of its stability in the presence of both peptidase and nuclease enzymes.[55] Various PNAs have been tested in the presence of mammalian and bacterial peptidases and been found to be stable.[56] In addition, in tests carried out using nuclease and DNase enzymes no degradation was observed.[57] For these reasons PNA is considered an attractive potential drug candidate.

It is thought that the reason why PNA is able to mimic DNA so effectively is due to the balance between the inherent flexibility of the molecule and the conformational constraint provided by the amide groups. This allows the molecule to adapt to variations in the structural features in the DNA target. It has been observed that any increase in the conformational freedom in aegPNA, by either elongation of the backbone or side chains, results in a loss of stability in PNA/DNA and PNA/RNA duplexes.[58]

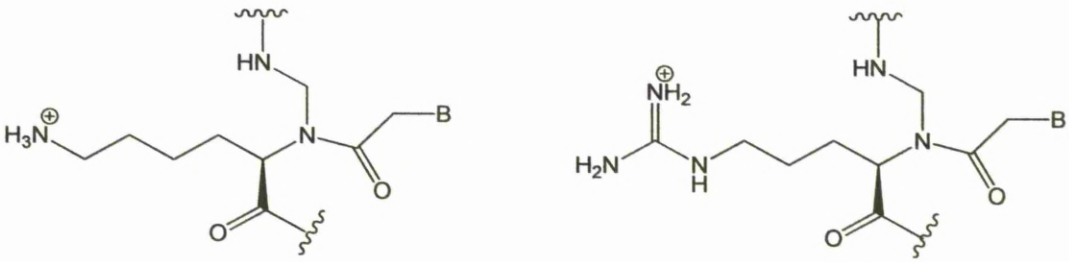


Figure 1.10: Modifications to the PNA backbone that increase aqueous solubility, (adapted from reference 51).[51]

The neutral nature of the pseudo peptide backbone also means that the electrostatic repulsion that is experienced between the two anionic strands in DNA and RNA does not occur in the PNA duplexes. While the neutral nature of the backbone might be an advantage, in terms of energy, it does bring about problems with aqueous solubility.[51] Common strategies that have been used to overcome its poor solubility include incorporation of hydrophilic amino acids, such as lysine (lysPNA), at the ends of the molecule,[59] incorporation of arginine residues which contain the guanidinium group (argPNA) into the backbone[60] and conjugation with cationic ligands[61] (Figure 1.10). All these methods have helped to improve solubility without compromising hydrogen bonding ability or specificity.

1.3.2 Secondary structures of PNA

During PNA duplex and triplex formation studies, it has been observed that the PNA strand can associate with its complementary oligonucleotide strand in either a parallel or antiparallel fashion, with the latter being the more stable. In comparison to an oligonucleotide sequence, the *C*-terminus of PNA is considered to be the 3'-end and the 5'-end the *N*-terminus. PNAs, like DNA and RNA, also follow the Watson and Crick rules for hydrogen bonding.[55]

PNA:DNA heteroduplexes are considerably more stable than the corresponding DNA duplexes and an increase in T_m of around 1K/base has been observed.[62] The helix tends to take on a form very similar to B-form DNA; the sugar pucker remains in the C2'-*endo* conformation and the helical rise remains the same. However, the helix does contain some elements of an A-form helix and some differences from the standard DNA helix form.[63] Whilst the helix is still right-handed[64] there are approximately 13 bases per helical turn in the PNA:DNA heteroduplex rather than the 10 seen in a standard DNA helix, the major groove is wider and, as a consequence of this, the minor groove is narrower and more shallow than A-form DNA. The amides of the PNA backbone are all orientated towards the solvent and in the *trans* conformation. It is interesting to note, however, that there is no observed hydrogen bond formation between the carbonyl and amide groups in the PNA backbone. The carbonyl groups within the nucleobase linker are all directed towards the *C*-teminus of the PNA strand.[62]

PNA:RNA duplexes are even more stable than PNA:DNA duplexes with a T_m increase of around 1.5K per base[65]. The PNA:RNA duplex is a right-handed helix with very similar properties to naturally occurring RNA, which usually takes on an A-form helix in which the sugar conformation adopted is C3'-*endo*. The amides of the PNA backbone switch geometry in this case to the *cis* conformation and again no hydrogen bonding is observed.[62]

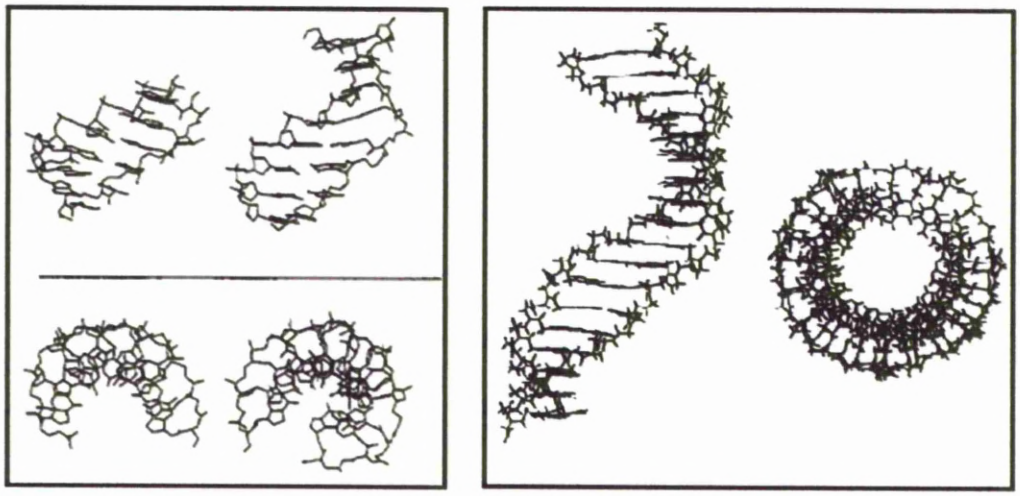


Figure 1.11: A PNA:RNA duplex and a PNA:DNA duplex (left) shown with side view (top) and a view down the helical axis (bottom) (taken from reference 68)[68] and a PNA:PNA duplex (right) again shown as a side view and a view down the helical axis (taken from reference 67).[67]

PNA can also hybridise with itself to form a PNA:PNA duplex which is a unique structure unobserved in nature; this has been termed "P-form" (Figure 1.11).[66] The two PNA strands associate through Watson-Crick interactions and produce a helix with a width of 28Å and 18 bases per turn, a lot more than the 11 seen in A-form DNA or 10 in B-form.[67] As a direct consequence of this, the helix has an extremely deep, wide major groove and shallow, narrow minor groove. Similarly to the PNA:DNA duplex, the backbone amide bonds are in the *trans* conformation and the amides of the nucleobase linker are orientated towards the *C*-teminus of the strands.[62]

Although originally designed as a $(DNA)_2$:PNA triplex forming strand, it has been found that this type of triplex only occurs in cases where a C-rich PNA strand and GC-rich DNA duplex are present.[46] For other single stranded sequences, the usual triplex structure is of the form $(PNA)_2$:DNA and $(PNA)_2$:RNA strands have also been observed,[69] both exhibit equally high stability. PNAs rich in pyrimidines form triplexes with double stranded DNA in the $(PNA)_2$:DNA format but do so by an unusual binding mode known as strand displacement or triplex invasion. One strand of the DNA duplex is displaced to form a loop and a triplex is formed inside the host

duplex, the PNA is bound to the DNA strand by both Watson-Crick and Hoogsteen base pairing and the triplex forms an unusual "P-helix" (Figure 1.12).[68]

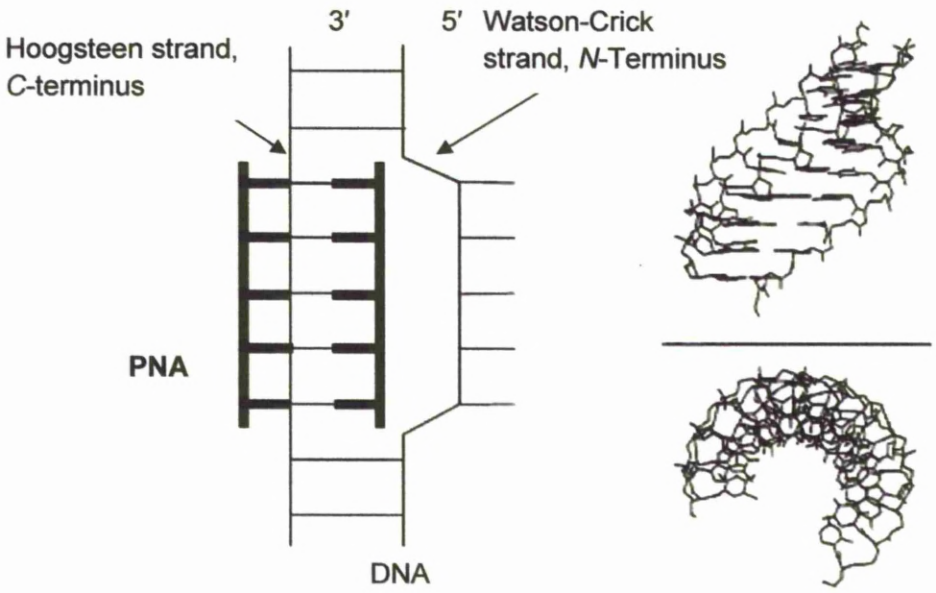


Figure1.12: A representation of strand displacement in a DNA duplex by a PNA duplex (left), (adapted from reference 68)[68] and a $(PNA)_2$:DNA P-helix (right) showing a side view (top) and view down the helical axis (bottom), (taken from reference 68).[68]

Although PNA forms excellent hybrid molecules with both DNA and RNA, the deviations away from the ideal B-form and A-form helix suggest that there is room for improvement. This has been attempted using various modifications of the PNA monomers in an effort to induce additional conformational constraint in the PNA backbone.[51] The most successful modifications have come from the introduction of cyclic motifs to the backbone. The highest rate of success has been demonstrated by the introduction of cyclopentyl (cypPNA)[70] and pyrrolidine (pyrPNA)[71] motifs (Figure 1.13). However, during research with these monomers, it has been found that the stereochemistry of such a constrained backbone, as well as other factors, do play an important role in host binding ability. The pyrrolidine derived analogues showed a distinct preference for antiparallel DNA binding over RNA or complementary binding,[71] and cyclopentyl analogues showed binding, with a slight gain of stability compared to aegPNA, only when in the *trans*- configuration,[72]

Significant work has already been carried out in the field of PNA structure and potential functions and interest in this type of molecule remains high today. Further research into perfecting the structure, increasing aqueous solubility and cell permeability may lead to some extremely important biological applications in the future.

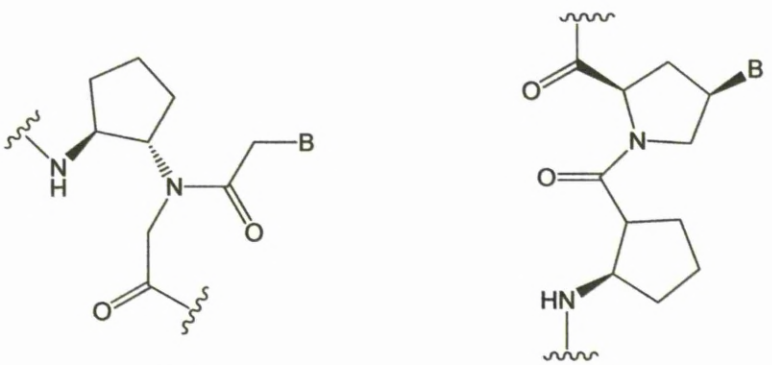


Figure 1.13: Modifications to the PNA backbone to induce conformational constraint.

1.4 α-Peptides

In nature, there are twenty commonly occurring amino acids which exhibit an α-arrangement of the functional groups, that is, the carboxylic acid and amine moiety are both attached to the same carbon atom. When condensed together, these form peptides and these peptides have the ability to hydrogen bond due to a high proportion of hydrogen bond donors and acceptors. This hydrogen bonding can be both intermolecular, which gives rise to β-sheet and associated structures, and intramolecular, which gives rise to α-helix structures.

The key to secondary structure in peptides is the planar peptide bond which was first published in 1951 by Pauling *et al.*[73] This was based on earlier work, in the 1930's, by Astbury[74] who proposed that two different forms of protein, α and β, existed in natural fibres. This proposal was based on a difference in x-ray diffraction patterns between stretched and none stretched wool fibres. Although the findings from this study were not accurate, they were the first major step in the examination of peptide secondary structure. The first to identify that the models proposed by Astbury were inaccurate was Neurath in a paper from 1940.[75] It was shown that the models contained clashes of atoms and were therefore not feasible.

The planar peptide bond model proposed by Pauling was based on resonance structures. It was suggested that all the atoms of the bond must be in the same plane in order for the entire structure to be stabilised by resonance (Figure 1.14).

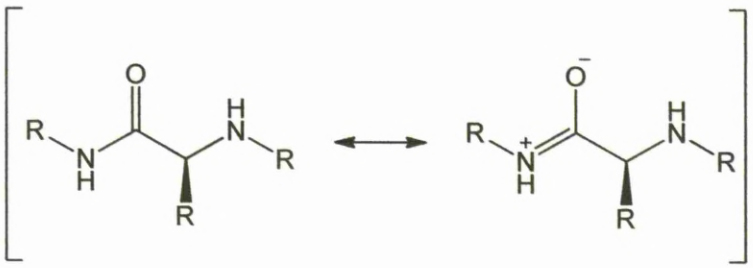


Figure 1.14: Amide bond resonance forms.

The planarity of the amide bond constrains the entire peptide chain which in turn means that there are only limited sites where intramolecular hydrogen bonding can occur. Based on this hypothesis, in collaboration with Corey and Branson,[73] two helical structures were proposed, modelled on the results of crystal structure studies. Pauling *et al.* constructed helices with planar amide groups and an observed hydrogen bond length of 2.72Å. They found there were only two possibilities, one with 3.7 residues per helical turn and one with 5.1; these were eventually termed an α-helix[76] and a γ-helix,[73] respectively. However, to this date, a γ-helix has never been observed (Figure 1.15).

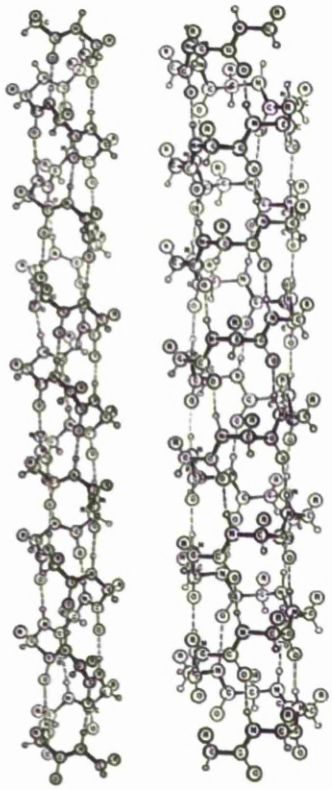


Figure 1.15: Pauling's α-helix (left) and γ-helix (right), (taken from reference 77).[77]

A major flaw in the work of Pauling *et al.* was the handedness of the helix. In the structure proposed, the helix was shown to be left-handed and made up of D-amino acids. We now know that the converse is true and α-helices are right-handed and consist of naturally occurring L-amino acids.[78] The modern day α-helix is also known to have a distance of 1.5Å between adjacent amino acids, 3.6 amino acid residues per turn and a pitch of 5.4Å. The hydrogen bonding in this type of helix exists in an [$i + 4 \rightarrow i$] fashion which denotes that the backbone N-H of one residue forms a hydrogen bond to the carbonyl of the amino acid 4 residues previous.[79] Variations of this type of hydrogen bonding in an α-helix exist but they are rare and make up only a small amount of amino acid structures. These include the 3_{10} helix[80, 81] (a tighter helix) and the π-helix[79, 82] (a looser helix).

The helix forming propensity of a peptide depends largely on the amino acid sequence present; some amino acids have a higher propensity for helix formation than others. In 1974, Chou *et al.* [83] derived a helix propensity scale based on observations from crystal structures which could be used to predict the folding properties of a peptide based only on the amino acid sequence. Structural prediction of this type has been an extremely active area of research over the past

30 years and more recent helix propensity scales allow α-helices to be predicted with up to 60% accuracy.[84, 85] It has been found that the amino acids with the highest helix forming propensity are alanine, uncharged glutamic acid, leucine and methionine whereas the lowest are proline, due to its lack of an amide proton, and glycine because of its high flexibility.[83] This supports earlier research which suggested that Ala, Leu and Glu were the most commonly found amino acid residues in helices.[84] Despite the immense amount of time and effort dedicated to the study of α-helices, the relationship between all the complex factors involved in folding is not yet fully understood.[86]

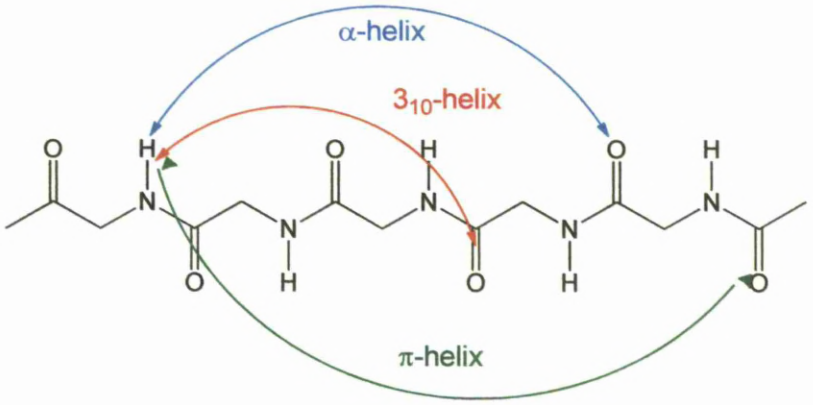


Figure 1.16: Hydrogen bonding patterns in different types of α-helix.

In the same year as the publication of the α-helix structures, Pauling *et al.* also published the first work on β-sheets.[87] The β-sheet secondary structure is only marginally less common than the α-helix and is associated with the formation of aggregates and fibrils observed in human diseases such as Alzheimer's.[88] This observation makes β-sheets an important area of research.

β-Sheets are networks of polypeptides (termed β-strands) joined together by intermolecular hydrogen bonds. The β-strands adopt a slightly extended conformation with a distance of 3.5Å between adjacent amino acids which is similar to a helix with only 2 residues per turn. The 'pleating' of the sheets is due to the sp^3 angle of the side chain bearing carbon causing the neighbouring carbon atoms to be either above or below the plane of the sheet(Figure 1.17).[89, 90]

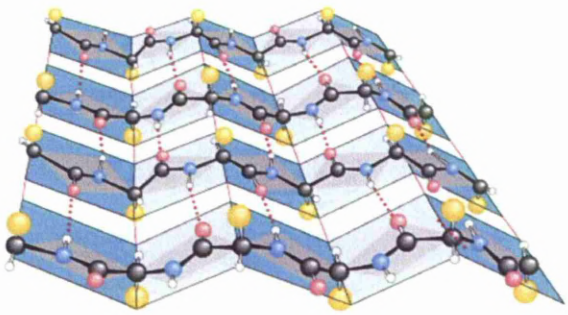


Figure 1.17: 'Pleating' in β-sheets (taken from reference 91).[91]

The two modern day β-sheet structures are termed parallel and anti-parallel[92] (Figure 1.18). The parallel form of the β-sheet is more pleated than its anti-parallel counterpart and is less commonly found. The parallel sheets are rarely found with less than 5 interacting strands ordered *C*-terminus to *C*-terminus and are uniform in nature with a consistent pattern of 12-membered hydrogen bonded rings. Anti-parallel β-sheets, on the other hand, can be found with as few as 2 interacting strands ordered in an *N*-terminus to *C*-terminus fashion and have a repeating pattern of 10-membered and 14-membered hydrogen bonded rings.[93]

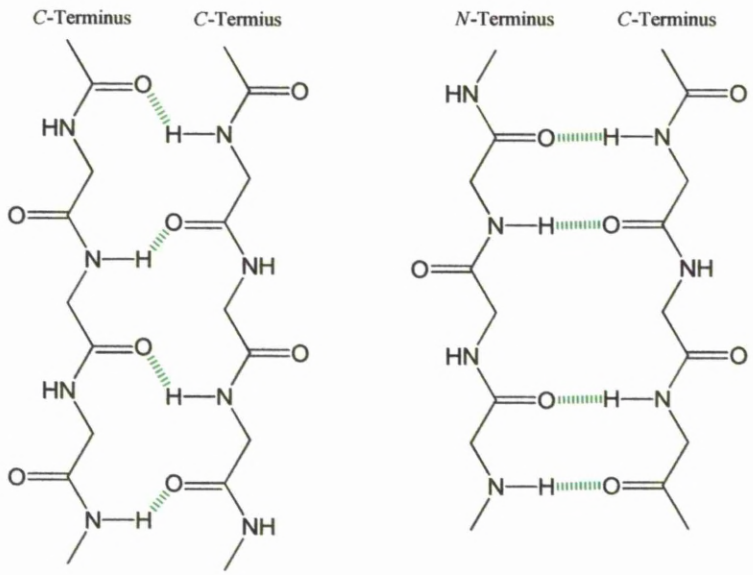


Figure 1.18: Hydrogen bonding in β-Sheets: parallel (left) and anti-parallel (right)

Naturally occurring macromolecules demonstrate such an array of complex secondary structures, which contribute to their important functions, that it is therefore of no surprise that research into mimics of these structural motifs is such a huge area. Unnatural mimetics of peptides and related structures which adopt well defined secondary structures have been termed 'foldamers'.[94]

1.5 Foldamers

Foldamers are described as 'a synthetic polymer with a strong tendancy to adopt a specific compact conformation'.[94] Over the past 20 years, research into foldamers has progressed rapidly and applications ranging from biomimetics[95] to nanotechnology[96] have been demonstrated. The range of structures and backbones which have been shown to adopt compact folded motifs is extremely diverse and constantly expanding with amino acids,[97] aromatic oligoamides,[98] carbohydrates[99] and polyisocyanates[100] included in the scope of this area. The following sections will concentrate on β-amino acids, carbohydrate and nucleic acid base-containing foldamers.

1.6 β-Peptides

1.6.1 Introduction

There have been countless studies into synthetic oligomers, but the most intensely studied are the β-peptides and their folding properties are the best understood. β-Peptides are made up of unnatural β-amino acids which have an additional methylene spacer between the amine and carboxylic acid functional groups compared to their corresponding α-derivatives (Figure 1.19).

Figure 1.19: Amino acids.

β-Amino acids are much less abundant in nature than their α-counterparts; the only natural β-amino acid commonly found is β-alanine found in the dipeptides carnosine and anserine.[101] β-Peptides have a number of attractive properties which make them good candidates for development into therapeutic agents. This will be discussed within this chapter.

The introduction of the extra methylene group in β-amino acids results in two possible sites for substitution, the C^2 and C^3 positions. Amino acids which are substituted on C^2, the carbon adjacent to the carbonyl carbon, are termed β^2-amino

acids and amino acids substituted on C^3, the same carbon that carries the amine, are termed β^3-amino acids. Most substituted β-amino acids prefer to reside in a gauche conformation around the C^2-C^3 bond, the one exception to this being a *syn*-C^2-C^3-disubstituted derivative which prefers an anti-periplanar conformation. In order for peptides to fold, the gauche conformation is required and as a result of this *syn*-C^2-C^3-disubstituted amino acids tend to form sheet- like structures.[102]

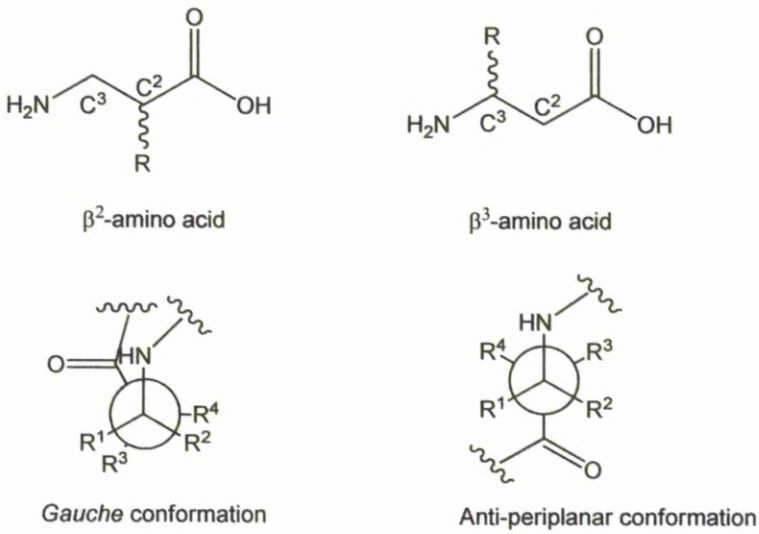


Figure 1.20: Substitution patterns in β-amino acids (top) and conformations about the C^2-C^3 bond (bottom).

The first to coin the term 'foldamer' was Gellman,[94] who has played a pivotal role in the determination of folding patterns in β-peptides. One of the main reasons for folding in α-peptides is the lack of nearest neighbour interactions leading to a preference for interactions between hydrogen bond donors and acceptors that are non-adjacent.[103] Early model studies by Gellman *et al.*, on both β- and γ-amino acid oligomers, showed that the latter formed a preference for nearest neighbour interactions whereas the former did not, suggesting that β-peptides showed more promise for folding studies.[104] Over almost three decades of work on β-peptide foldamers, a large array of structures have been observed including pleated sheets,[105] turns,[106] helices[107] and more. The unnatural peptides experience much more structural diversity within these classes than their natural analogues e.g. β-peptides can from helices with hydrogen bonded rings containing between 8 and 14 members (Figure 1.21) whereas the dominant form in natural α-helices is the 13 membered hydrogen bonded ring.

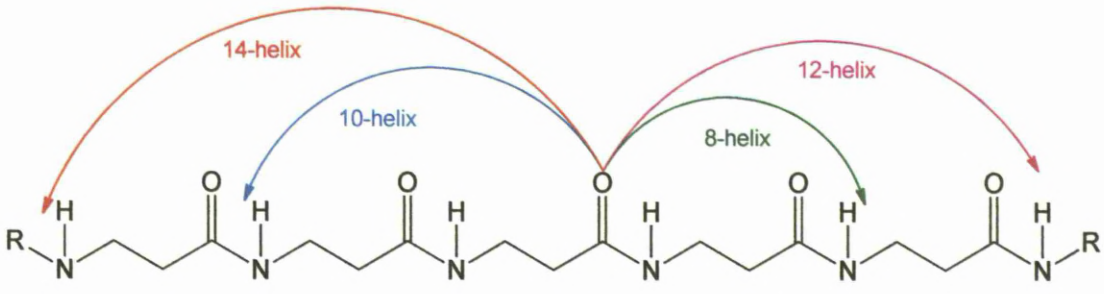


Figure 1.21: Helices found in β-peptides.

Early research by Seebach *et al.* into naturally occurring polyhydroxybutyrates (PHB's, Figure 1.22) and their synthetic analogues showed that these structures adopted a 14-helix with 3 residues per hydrogen bonded turn (3_1-helix) and that replacement of the backbone oxygens with nitrogen was highly tolerated.[108] This led to experiments using the β-analogues of valine, lysine and leucine assembled into a β-hexapeptide of the form (H-β-HVal-βHAla-β-HLeu)$_2$-OH. It was found that this hexapeptide also formed a 14-helix with 3 amino acid residues per turn. However, although circular dichroism spectra showed similarities between the hexapeptide and α-peptides which adopt a β-sheet conformation, NMR studies in d_5-pyridine revealed that the hydrogen bonds were in fact intra- not intermolecular.[109, 110] For comparison, peptides made up from the naturally occurring α-amino acids form a 13-helix with 3.6 amino acid residues per hydrogen bonded turn (3.6_1-helix).The major difference between natural and unnatural peptides is the direction of the helix; the unnatural helix dipole runs towards the *C*-terminus, the opposite of what is observed in the natural helix.[111]

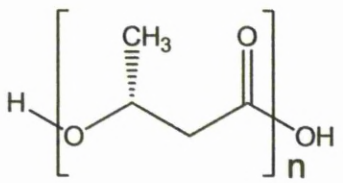


Figure 1.22: General structure of a polyhydroxybutyrate.

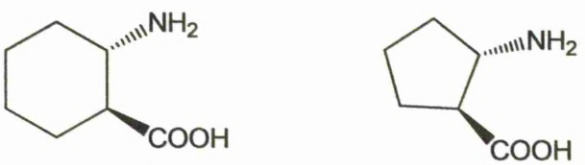
This was considered a surprising result due to the fact that, for an α-peptide to form a stable secondary structure in solution, it must contain between 15-20 residues whereas the unnatural peptides exhibited secondary structure with only 6.[112] It had been expected that the extra carbon atom in the β-amino acid structure would provide the peptide with additional flexibility and, thus, it had been proposed that a

lot more residues would be required in order for these peptide molecules to form ordered secondary structures.

1.6.2 Conformationally-constrained β-peptides

The expectation of increased flexibility is reflected in the work of Gellman *et al.* conducted at around the same time on conformationally-constrained β-peptides.[114] Previous studies on β-alanine peptides had shown that this increased flexibility and lack of side chains caused the oligomer to be unordered in solution, although it formed β-sheet motifs in the solid state.[113] Gellman postulated that reducing the number of degrees of freedom available to the peptide would help to induce helical folding. This was achieved by incorporating the C^2 and C^3 carbons of the backbone into a cyclohexane ring and utilising a *trans*-arrangement of the functional groups(Figure 1.23), thus trapping the amino acid in its gauche conformation. This design was first investigated by computer modelling which indicated that a stable helix could be derived from a 10 residue oligomer and that this arrangement formed a 14-helix. This was subsequently investigated experimentally and oligomers of *trans*-aminocyclohexanecarboxylic acid (*trans*-ACHC, Figure 1.23) of various lengths were synthesised. A crystal structure of the hexamer revealed that the computer modelling was in fact accurate. These constrained oligomers did indeed form a 14-helix that was shown to be extremely stable in solution.[114]

Based on this result, it was postulated that a reduction in ring size should result in a smaller hydrogen bonded ring and hence a tighter helix. This hypothesis was investigated using a hexamer of *trans*-aminocyclopentenecarboxylic acid (*trans*-ACPC, Figure 1.23). Interestingly, it was found from 2D-NMR studies that this oligomer did indeed form a tighter 12-helix which was stable in organic solvents.[107,115]



Trans-ACHC *Trans*-ACPC

Figure 1.23: Conformationally constrained amino acids developed by Gellman.

Since the discovery that β-peptide secondary structure formation could be controlled by relatively small changes to the backbone, there have been countless studies into backbone variation. Examples of these include oxetane,[116] cyclobutane[117] and

cyclopropane[118] motifs. The first to investigate 4-membered cyclic amino acid monomers was Fleet and co-workers. They found through 2D-NMR studies that a hexamer compound of a *cis*-oxetane-β-amino acid monomer (Figure 1.24) provided a left-handed 10-helix, when present as a solution in either chloroform or benzene.[116] More recently, studies have shown that the corresponding oligomer made from the carbocyclic derivative, *cis*-aminocyclobutanecarboxylic acid (*cis*-ACBC, Figure 1.24), exhibited only nearest neighbour interactions and thus did not adopt a helical conformation.[119] However, changing the relative positions of the functional groups from *cis*- to *trans*- induced a conformational change, presumably to the gauche conformation. Dipeptides of *trans*-ACBC and mixed ACBC showed a preference for an 8-helix structure[120] whereas hexamers and octamers of *trans*-ACBC showed a preference for longer range hydrogen bonding and the formation of a 12-helix in d_5-pyridine was observed. [117]

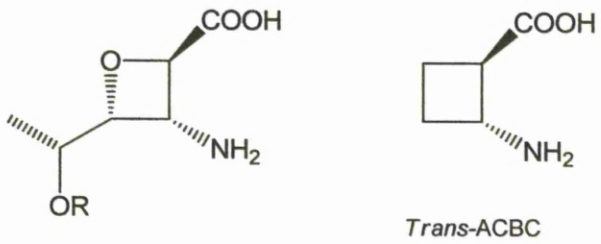


Figure 1.24: Oxetane (left) and cyclobutane (right) derived β-amino acids.

Seebach *et al.* have shown that when achiral cyclopropane amino acids (Figure 1.25) are assembled into peptides, they form a ribbon like arrangement of 8-membered hydrogen bonded ring.[118] This deviation away from the traditional conformational constraint about the C^2-C^3 bond has shown just how adaptable these peptides are. It is also interesting to note that when the 3-membered cyclopropane ring was replaced with a cyclohexane (Figure 1.25) the shape adjusted from an 8-helix to a 10-helix.[121] 8-Helices have been observed in a number of oligomers and are thought to be the result of steric and electronic effects which favour nearest neighbour type interactions over longer range interactions.

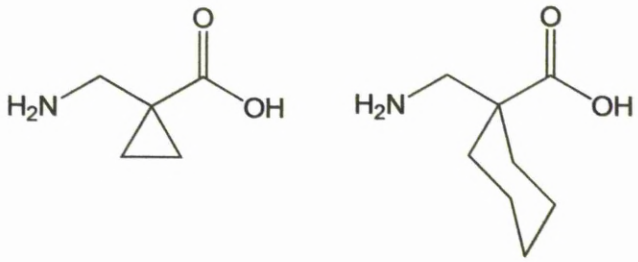


Figure 1.25: Cyclopropane (left) and cyclohexane derived β-amino acids

1.6.3 Mixed helix forming β-peptides

It has been observed that while homo-oligomers tend to adopt a highly ordered regular helix, oligomers with mixed residues can form what is termed a 'mixed helix'. The most well known example of this is the 10/12 helix. β-Peptides with an alternating β^2- and β^3-substitution pattern are known to adopt this type of conformation with the amides attached to the unsubstituted methylene groups hydrogen bonding to each other forming a 10-membered ring and the amides attached to the carbons bearing side chains adopting 12-membered rings (Figure 1.26).[122] These structures have been studied by 2D-NMR and circular dichroism in both pyridine and methanol and it has been found that there are two types of amide bond alignments in this type of helix. The 10-atom ring amides are orientated almost perpendicular to the helical axis and the 12 atom ring amides are orientated with the helical axis. This results in a reduced helix dipole compared to the uniformly aligned 12 and 14 helices.[122]

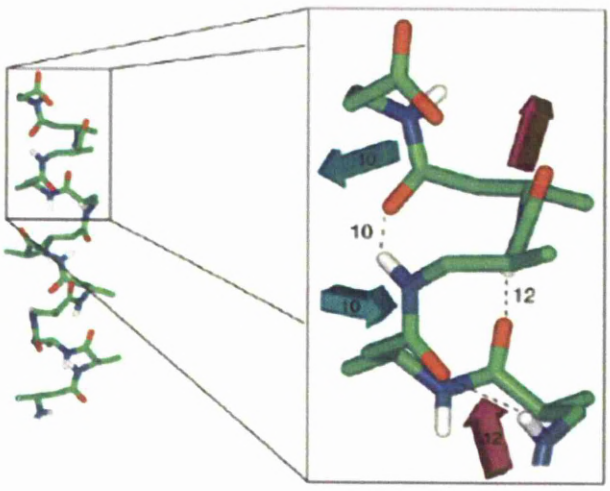


Figure 1.26: Structure of the 10/12-helix with an expanded view showing orientation of 10-atom ring amides (green) and 12-atom ring amides (purple), (taken from reference 102)[102].

It has been postulated that the 10/12-helix is favoured when residues with intercalating side chains are placed 3 residues apart and this has encouraged computational modelling of these types of oligomer. Applying these techniques to oligomers of β-alanine predicted that 10/12-helix formation was favoured over a 14-helix.[123] Introduction of a single methyl group at either the β^2 or β^3 positions allowed for six patterns of dipeptide repeat, (*S*)-β^2/(*S*)-β^2, (*S*)-β^2/(*R*)-β^2, (*S*)-β^2/(*S*)-β^3, (*S*)-β^2/(*R*)-β^3, (*S*)-β^3/(*S*)-β^3 and (*S*)-β^3/(*R*)-β^3 and each of these were modelled. Results of the study indicated that there seemed to be no additional stabilisation effect to the 10/12 helix but rather a destabilising effect on the unfavoured conformations, in most cases the 14-helix.[123] The predictions gained from this study are yet to be tested experimentally but could provide a useful tool in the design of specific helices.

1.6.4 Problems with β-peptides

A common problem encountered with β-peptides is with their aqueous solubility[107] and a number of methods have been devised to tackle this issue. One of the first strategies devised to overcome this problem involved simply adding a charged component to the peptide backbone. Gellman *et al.* were the first to attempt this strategy by adding amino groups in an alternating fashion to the cylcohexane moiety of the *trans*-ACHC hexamer mentioned earlier (Figure 1.27).[124] This hexamer could

bear a charge of +3 at pH ≤7 and indeed did have good aqueous solubility. However, it produced poor resolution in NMR studies.

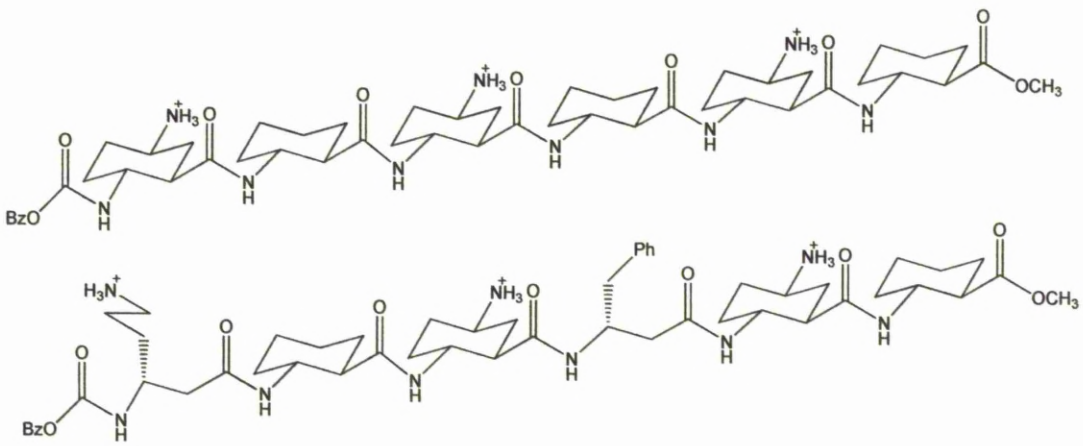
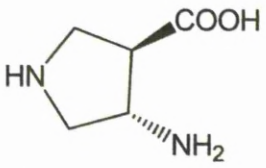


Figure 1.27: Gellman's hexamer with alternating charges (top) and hexamer with 2:1 cyclic to acyclic residues (bottom).

A hexamer with a 2:1 cyclic to acyclic residue ratio produced much better resolution in NMR studies in an aqueous acetate buffer at pH 3.9 and the long range coupling constants indicated that it did indeed form a 14-helix in aqueous solution.[124] This was supported by a characteristic CD maximum at 215nm, which is comparable to the *trans*-ACHC CD spectral findings. The CD spectrum for the fully cyclic charged hexamer also produced a strong maximum at 215nm which indicated 14-helix formation, although without the corresponding NMR data this could not be confirmed for certain.[124]

A second strategy developed to overcome this problem was to replace the *trans*-ACPC carbocycle with the charged pyrrolidine analogue *trans*-3-aminopyrrolidine-4-carboxylic acid (*trans*-APC, Figure 1.28).[125] This both retained the structural rigidity of the original oligomer and improved aqueous solubility as it formed a helix which was ampiphilic in nature. This type of oligomer has been shown to have excellent antibacterial properties against a number of bacteria strains and also excellent selectivity for bacterial cells over mammalian.[125]



Trans-APC

Figure 1.28: *Trans*-aminopyrrolidinecarboxylic acid

1.6.5 Sheet structures

As well as forming helices, β-peptides can also mimic another important class of biologically active motifs, namely β-sheets. These unnatural peptides can adopt one of two types of anti-parallel arrangements, whereas the natural α-peptides that can only form one. This variation exists because of the ability of β-peptides to reside in two conformations about the C^2-C^3 torsion angle, 'anti' and 'gauche'. When all the residues in the peptide adopt the 'anti' conformation, the sheet is said to be type I. This type of sheet results in a net dipole due to the parallel arrangement of the carbonyl groups. When all the residues of the peptide adopt the 'gauche' conformation, the sheet is said to be type II. This sheet has little or no net dipole due to the anti-parallel arrangement of the carbonyl groups; this is comparable to the naturally occurring α-peptides which adopt a β-sheet motif.[126]

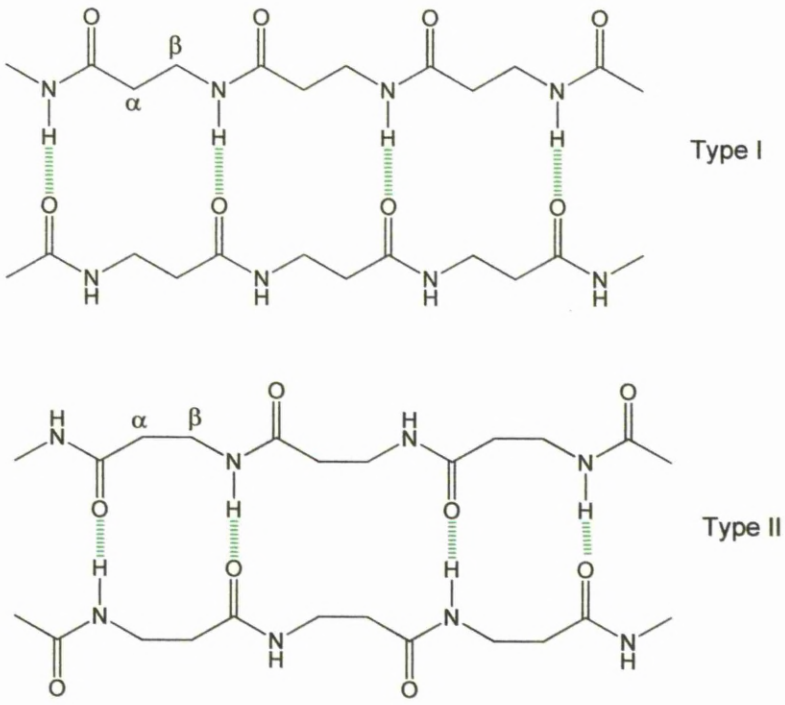


Figure 1.29: Two types of β-sheet structures observed in β-peptides.

As mentioned previously, early work on poly-β-alanine showed that this peptide tended to form sheet structures in its crystalline form, but was relatively disordered in solution.[113] Sheet secondary structures have also been observed in other poly-β-amino acids. Work by Gellman *et al.* has provided an important insight into the preparation of β-peptides with sheet-like conformations. A hybrid molecule containing two C^2, C^3-disubstituted amino acid residues was shown to adopt a hairpin-like, type I anti-parallel sheet conformation both in organic solvents and the solid state. The molecule contained a central L-proline-glycolic acid segment which assisted the β-turn conformation and initiated the formation of an anti-parallel sheet containing both 10-membered and 14-membered hydrogen bonded rings (Figure 1.30, **A**).[126] One of the key structural features of the two β-amino acids required for generation of this sheet conformation, is the *syn* arrangement of the C^2 and C^3 substituents. This promoted an 'anti' torsion angle about the C^2-C^3 bond. Replacement of the disubstituted residues with β-alanine led to the formation of a type II non-polar sheet while incorporation of two β^3-substituted amino acids induced an equilibrium between the two types of β-sheet. This research suggested that β-amino acids with a *syn*- C^2, C^3 disubstitution pattern would have the highest propensity for β-sheet formation.

Further to this finding, a similar β-sheet motif has been synthesised from entirely β-amino acid residues by replacing the L-proline-glycolic acid segment of **A** with a heterochiral nipecotic acid dipeptide (Figure 1.30, **B**). The conformational constraint of the heterochiral nipecotic dipeptide fragment induced β-sheet formation. However, it was found that if a homochiral dipeptide (i.e. having the same configuration at carbons **1** and **2** in Figure 1.30, **B**) was used instead interaction of the two terminal amino acid residues was inhibited, thus preventing sheet formation.[127]

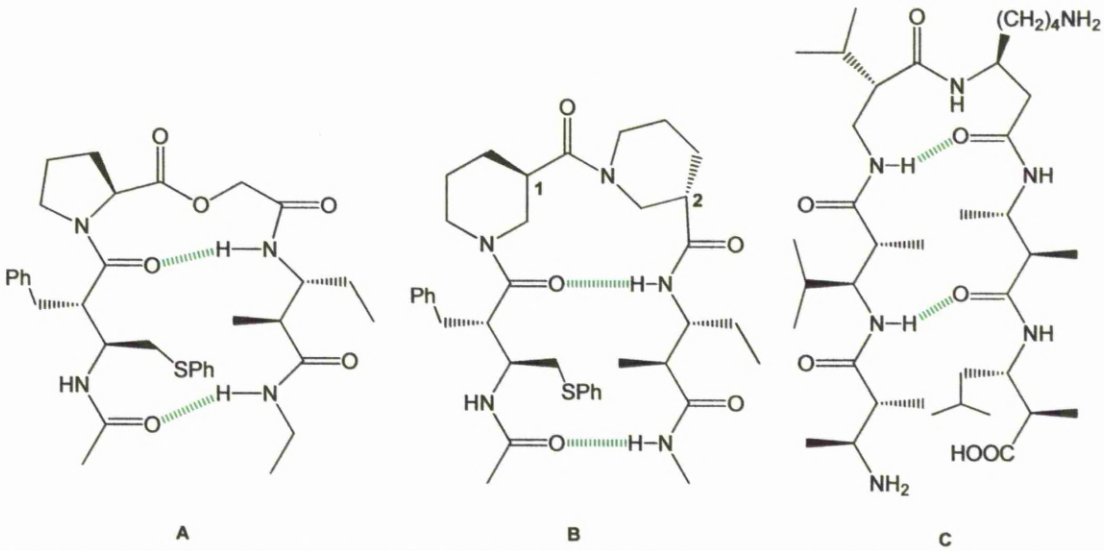


Figure 1.30: β-Peptide hairpin type structures by Gellman (**A** and **B**) and Seebach (**C**)

Research by Seebach *et al.* provided more evidence for the formation of β-peptide hairpins in organic solvents by preparing hexapeptides in which the first and last two amino acid residues were *syn*- C^2, C^3 disubstituted (Figure 1.30, **C**).[128] Interestingly, this group reported that a turn-inducing dipeptide containing a C^2- followed by a C^3-substituted β-amino acid provided a different reverse turn motif to that observed in the work of Gellman with the nipecotic acid dipeptides. A 10-membered hydrogen bonded ring was observed by Seebach whereas Gellman's peptide gave a 12-membered hydrogen bonded ring. This discovery, once again, highlights the greater conformational diversity of the β-peptides as opposed to their natural congeners, in which only one type of reverse turn (β-turn) is commonly observed.[102]

More recently, it has been found that conformational control can be exerted on certain peptides whilst in aqueous solution by using metal cations such as zinc. This type of conformational control is dependent on the nature of the amino acid residues and where they are placed in the in the peptide chain. A β-octapeptide has been

synthesised with terminal β-histidine and β-cysteine residues.[129] These residues show a high affinity for zinc cations, through interaction with the imidazole *N*-atoms in the case of histidine and the thiol in the case of cysteine. The peptide was examined by NMR in methanol (with *S*-protection of the terminal β-cysteine residue to prevent any disulphide formation) and it was found to form a 14-helix. Upon deprotection and isolation with $ZnCl_2$ in an aqueous environment, the peptide was seen to undergo a switch from a helical structure to a hairpin-like structure containing a 10- and two 14-membered hydrogen bonded rings. It is believed that interaction of the two terminal amino acid residues with the zinc cation induced and stabilised this change.[129] This research highlights new avenues and applications for β-peptides which could prove to be important in the future.

1.6.6 Cyclic β-peptides

As well as linear β-peptides, studies have also been carried out on a number of cyclic β-peptides. Various groups have undertaken research on these types of peptides and have shown that they can undergo molecular stacking to form column type structures through hydrogen bonding interactions between the amine group of one macrocycle with the carbonyl groups of another. They can also exhibit sheet-like structures through different hydrogen bonding interactions.

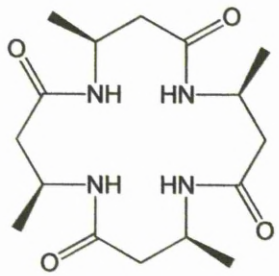


Figure 1.31: Cyclic tetrapeptide

One of the first investigations into cyclic β-peptides was carried out by Seebach *et al.* in which a series of methyl-substituted cyclic tetrapeptides (Figure 1.31) were synthesised and studied by powder diffraction. The results suggested that the rings did indeed stack on top of each other in an ordered fashion and were stabilised by hydrogen bonding interactions between the rings.[130]

More recently, conformationally constrained cyclic tripeptides have been synthesised. It has been found that these also exhibit the same type of

intermolecular hydrogen bonding to give tubular structures when crystallised from a mixture of formic acid and water (Figure 1.32).[131] Interestingly however, it was found that when the concentration of the peptide was increased, rather than crystals, a fibrous material was observed. This was later confirmed to be a sheet-like structure by electron diffraction analysis.

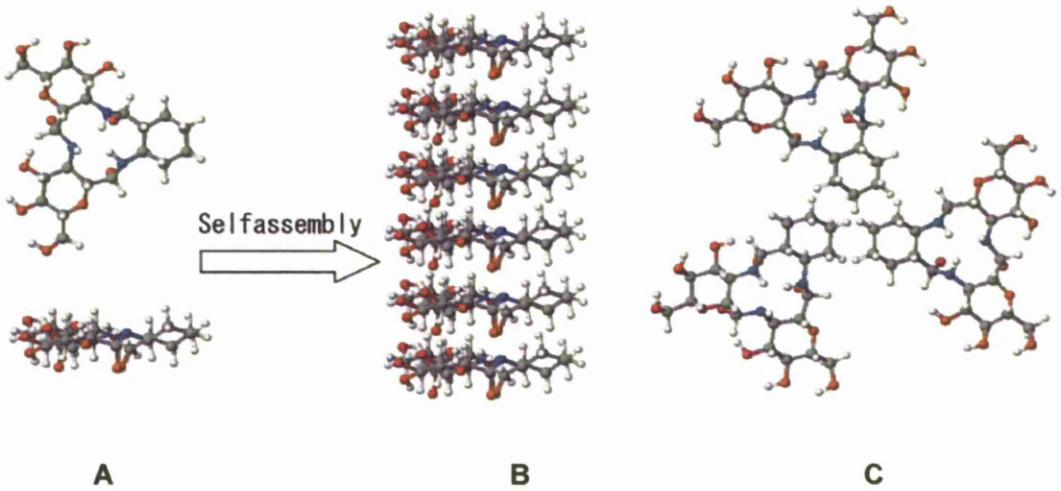


Figure 1.32: Cyclic β-tripeptides (**A**) assembled into tubular form (**B**) and sheet form (**C**) (taken from reference 131).[131]

Further to this, Seebach also demonstrated that cyclic β-peptides may also provide good synthetic analogues of naturally occurring hormones. This group synthesised a mimic of the hormone somatostatin which has many important biological functions, one of which is the regulation and release of growth hormone.[132] A synthetic analogue of this hormone (octreotide, Figure 1.33) already exists and is used therapeutically to treat certain cancers.[133] However, this analogue is made of α-amino acids and this results in it having an extremely short half life *in vivo* of just 90 minutes.

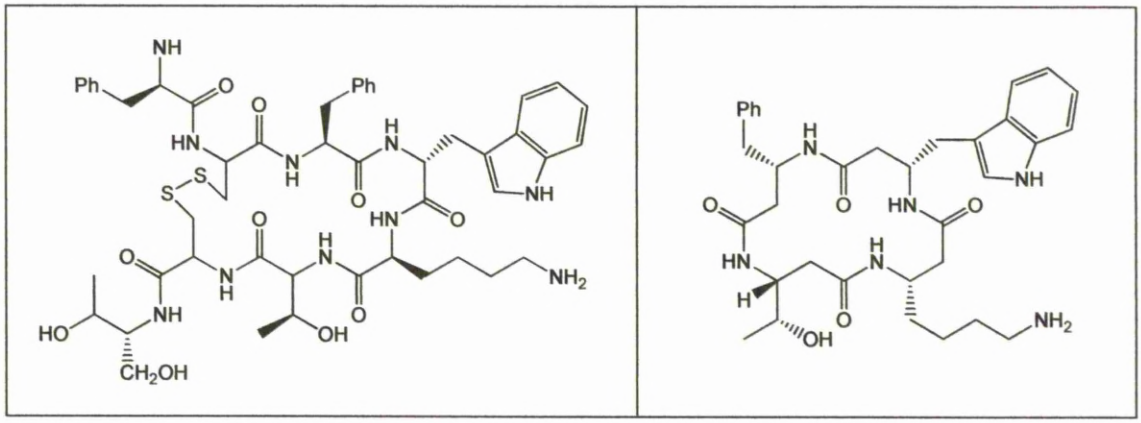


Figure 1.33: Octreotide (left) and a β-peptide mimic of Somatostatin (right).

The mimic synthesised by Seebach *at al.* (Figure 1.33) was shown to bind to various somatostatin receptors with affinity comparable to somatostatin itself. Subsequent NMR studies of both the existing synthetic hormone and this analogue showed significant, although not perfect, spacial overlap of the side chains which are essential for biological activity.[134] The results of this study indicate that, with further research, there may be a potential for β-peptides to be used as therapeutic agents within the area of artificial hormones.

1.6.7 Higher order structures in β-peptides

An ultimate goal of the study of β-peptides is to enable the construction of complex higher order structures, particularly artificial proteins. Significant progress has been made towards this goal with several groups publishing promising results. Conformational order in complex structures is a requirement in order to develop foldamers with sophisticated biological functions such as catalysis. Research by Gellman *et al.* has identified helix bundle motifs in peptides that contain alternating α- and β- amino acid residues and also in what they term a 'block' strategy which contains a mixture of both α-peptide helices and α/β-peptide helices which associate together to form a helix bundle.[135,136]

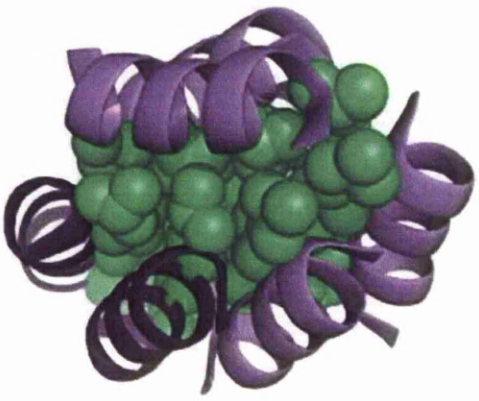


Figure 1.34: A synthetic β-peptide octamer (peptide helices shown in purple) that displays a hydrophobic core (shown in green) (taken from reference 138).[138]

Work by Schepartz *et al.* indicated that helices containing only β^3-amino acid residues can also assemble in this way. They observed that a 12-residue helix forming foldamer self-associated in aqueous solution, to from an octameric bundle held together only by non-covalent inter-residue interactions. This bundle also displayed similar properties to those associated with an α-helix bundle such as a hydrophobic core (Figure 1.34). [137,138] These studies indicate that, clearly, there is potential for synthetic mimics of natural biopolymers that may, in the future, provide some extremely desirable functions.

1.7 Carbohydrate containing foldamers

Carbohydrates have many important roles in biology ranging from energy storage to featuring as essential units in glycoproteins and hormones. They have, therefore, featured in much chemical and biological research, one of which is in the area of foldamers. There is a vast number of examples of foldamers that contain carbohydrate derivatives ranging from the 4 membered oxetanes,[116] mentioned previously, to pyranose-based carbohydrate analogues. In particular, carbohydrate amino acid derivatives have been shown to play a useful role in the construction of peptiomimetics,[139] the assembly of combinatorial libraries[140] and they have also produced interesting biological results when incorporated into peptides.[141]

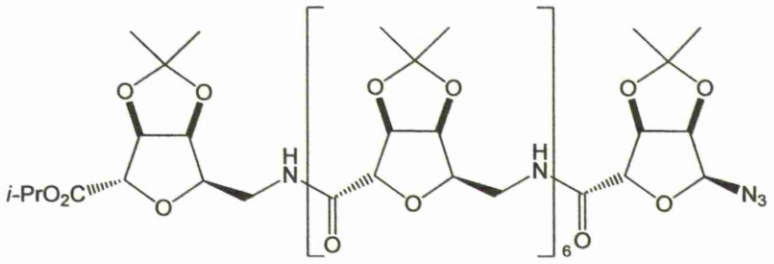


Figure 1.35: Furanose-based oligomer that forms a 16-helix.

Work by Fleet *et al.* has shown that oligomers of furanose dervied amino acids can form highly ordered structures in solution (Figure 1.35). An octamer of *C*-glycosyl α-D-lyxofuranose amino acids was synthesised and it was shown through NMR studies, to adopt a strongly helical structure in solution; a 16-helix.[142] Interestingly, shorter oligomers were shown to be disordered in solution which suggests that, in this case, more residues are required to induce a secondary structure.

It has also been shown that oligomers utilising pyranose carbohydrate amino acids can exhibit secondary structures in solution with as little as 4 residues. However, the secondary structure is stabilised by elongation of the chains. Work by Gervay *et al.* demonstrated the first pyranose carbohydrate oligomer containing amide bonds for which secondary structure was observed.[143] However, full elucidation of the interactions involved and conformation of the secondary structure was not published until 2 years later. This group found that the secondary structure appeared to vary with oligomer length; for the shorter oligomers, for example a tetramer, a 16-helix was observed whereas with longer oligomers, an octamer, a 22-helix was observed.[144] A promising property of these oligomers is that they are water soluble and seemed to keep their complex secondary structure when in aqueous solution.

The area of β-peptides derived from carbohydrate β-amino acids has also received much attention and numerous examples of these oligomers exist.[145] An interesting example of this comes from a recent paper by Andreini *et al.* in which a β-peptide with spirocyclic disubstitution and *N,O*-acetal functionality at the β-carbon of the backbone is presented. Studies by NMR and CD in deuterated DMSO indicated that even short oligomers of 4 residues in length adopt secondary structure and there was strong evidence for the formation of 8-membered hydrogen bonded rings.[146]

This research has been, and still is, extremely informative and important in terms of examining the possibilities of synthetic peptidomimetics. Carbohydrates have been placed in a position of particular importance due to their contribution to the ridigity

and stability of secondary structures, which are not only limited to helices but β-turns[147] and others. One particular interesting feature is the fact that the polyhydroxylated backbone offers the opportunity for custom synthesis of hydrophobic or hydrophilic derivatives, which is known to be of particular importance in protein tertiary structure assembly.[142]

1.8 Nucleobase-containing foldamers

Oligomers that combine the folding properties of peptides and specific recognition properties of nucleic acids have also have been of interest over recent years. Although most of the work into these types of foldamers has focused on PNA, there are a few notable examples of true polypeptides functionalised with nucleobases. The first example of a nucleobase-functionalised amino acid dates back to 1969 when Doel *et al.* synthesised a series of adenine, thymine and cytosine α-amino acids.[148] However, had these amino acids been sequenced into an oligomer, it is unlikely that association with a complementary oligonucleotide strand would have been successful due to reduced spacial distance between the nucleobases in the peptide.The most successful oligonucleotide mimics are those with a backbone isosteric to that observed in a natural oligonucleotide strand.

One of the first examples of a successful oligonucleotide binding nucleopeptide was reported in 1997 by Yamazaki *et al.*. This group synthesised a tetrapeptide made up of L-α-amino acids functionalised with adenine alternately spaced with glycine, serine, threonine or tyrosine. They found that the peptide, although short, bound to poly d(T) and poly d(U) sequences with a binding affinity higher than that found with a natural d(A) tetranucleotide. It is thought that the increase in binding affinity was due to a lack of electrostatic repulsion between the neutral peptide and negatively charged oligonucleotide backbones.[149]

Diederichsen *et al.* demonstrated that combining the helix-forming properties of β-peptides and association properties of nucleobases can lead to interesting results.[150] Using the knowledge that β^3-substituted peptides tend to form a 14-helix with 3 residues per turn, this group assembled peptides with a nucleobase side chain substitution every 3 residues.

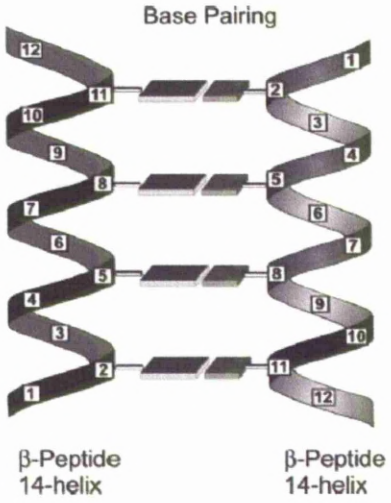


Figure 1.36: A model of base pairing in an anti-parallel β-peptide duplex functionalised with nucleobase side chains every 3 residues. (taken from reference 151).[151]

This substitution pattern resulted in one side of the helix being occupied by only nucleobases (Figure 1.36). [150] When in solution, anti-parallel duplex formation through the Watson-Crick base pairing of complementary sequences was observed for A-T, T-T, C-C^+, T-C, T-G pairs.[151] Although GCGC sequences were self recognising and displayed the strongest binding affinity of all base pairs, homo(G) and homo (C) sequences displayed no specific interactions.[152]

Previous work within our group has shown that oligomers made up of a thymidine derived β-amino acid form an 8-helix in solution.[153] Oligomers of 4 residues in length were studied by NMR and, when in a solution of d_5-pyridine, were found to assemble into a helix in which the thymine nucleobases were located on the outside of the helical chain. An overall 'zig zag' arrangement of the carbocyclic rings in the chain resulted in a pattern of alternating nucleobases being found on alternating 'sides' of the helix (Figure 1.37). [153] This result indicates that, in the future, peptides of this nature may be able to associate through base pairing to their complementary strands in a manner similar to that reported by Diederichsen.

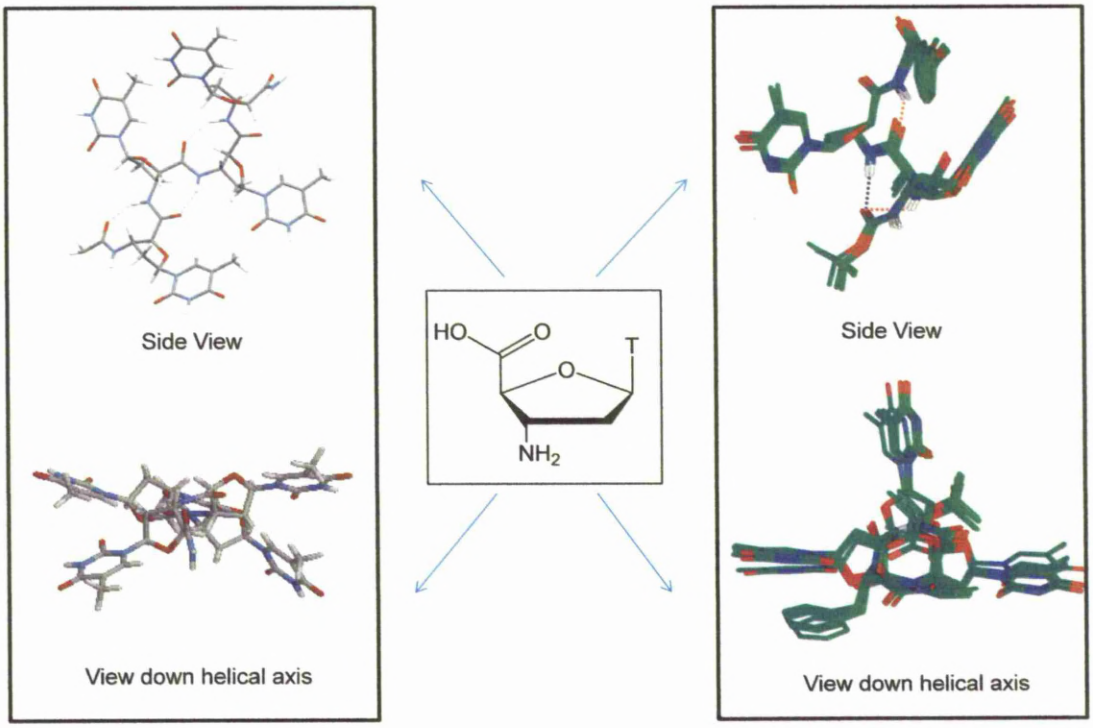


Figure 1.37: A thymidine derived β-amino acid (centre) and the oligomers assembled by our group (left) and the oligomers assembled by Chandrasekhar *et al.* (taken from references 153, 154 and 155)[155]

Further to this, Chandrasekhar *et al.* have shown that oligomers made up of the same thymidine-derived β-amino acid alternated with α-amino acids also formed a helical secondary structure. A mixed helix of an 11/8 nature was formed by these molecules and, similar to the work of our group, the nucleobases were all still found on the outside of the helix. However, this time the nucleobases were located in a 'radial' fashion around the helical axis rather than on alternating 'sides'.[154] Although not yet perfected, evidence from these and other studies point to nucleobase association becoming a useful tool in molecular design in the future.

1.9 Project aims

The primary goal of this project was to synthesise nucleoside β-amino acids from their natural nucleoside starting materials and assemble them into peptides. It was hoped that combining the helix forming properties of cyclic β-amino acids with the characteristics of nucleobases would provide novel foldamers with specific

recognition properties. Monomers were designed with protecting groups which were compatible with both solid and solution phase peptide synthesis so that both methods of oligomer assembly could be investigated. Work in this thesis concentrates on the synthesis of thymidine, 5-methyl-2′-deoxycytosine and 2′-deoxyadenosine amino acid derivatives and their assembly into peptides.

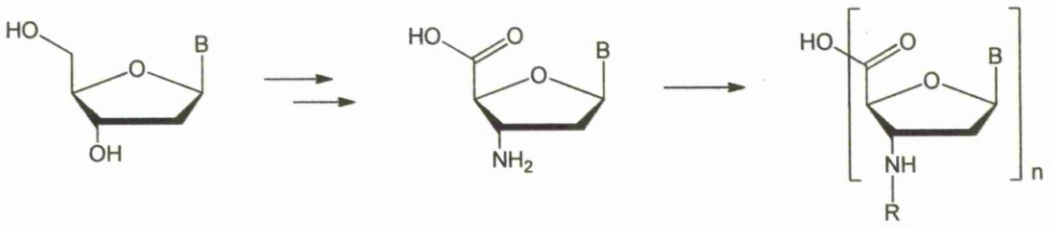


Figure 1.38: Basic synthesis of an oligomer, B=nucleobase.

The peptides were specifically designed to posses both a 5′- carboxylic acid and 3′-amine terminus so that the possible formation of cyclic structures that may associate through base stacking interactions could be explored.

Chapter 2: Results and Discussion 1

Synthesis of Nucleoside Derived β-Amino Acid Monomers

2.1 Proposed route for the synthesis of nucleoside β-amino acids

The overall planned synthesis of the required modified nucleoside monomer began with the parent nucleosides. The synthesis proceeded through several steps which commenced with protection of the 5′-hydroxyl group and any functionality included in the nucleobase. This was followed by a double inversion of stereochemistry at the 3′-position to introduce the azide functionality with overall retention of configuration. In general, all the synthetic routes involved reduction of the azide to form the amine and oxidation of the 5′-hydroxyl group to form the carboxylic acid (Figure 2.1). Finally, protection of either the amine or acid functionality provided a monomer which could be used for either coupling in solution or on a solid-phase resin. In particular, the synthetic work reported herein has focussed on β-amino acids associated with the nucleobases 5-methyl cytosine, thymine and adenine (Figure 2.1) and these will be addressed in turn.

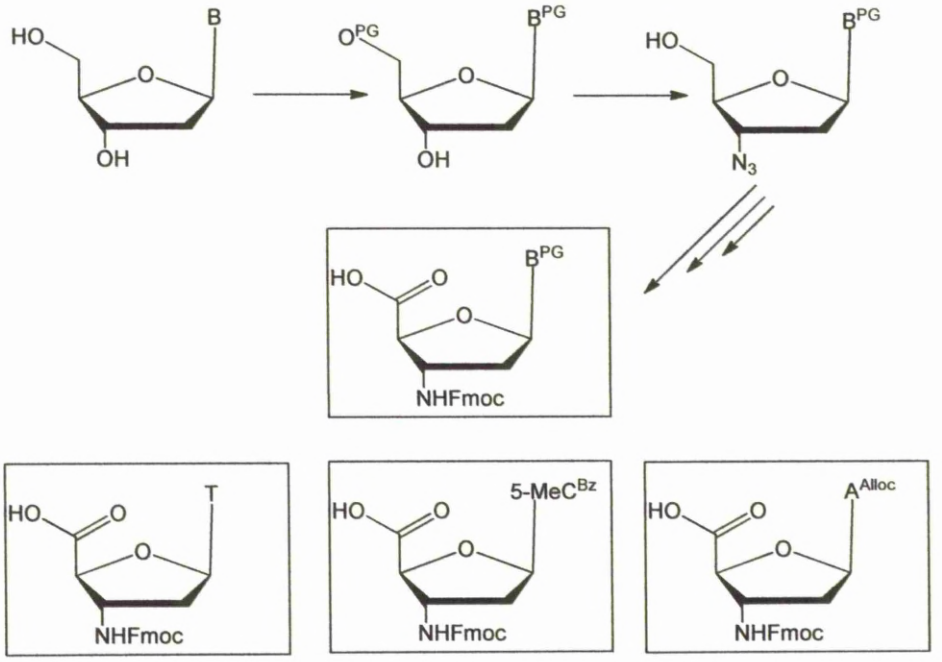


Figure 2.1: General reaction scheme outlining the synthesis of a nucleoside β-amino acid (B=nucleobase and PG=protecting group) and also showing the target β-amino acids (T = thymine, 5-MeC = cytidine, A = adenine, Bz = benzoyl, Alloc = allyloxycarbonyl and Fmoc = 9-fluorenylmethoxycarbonyl)

2.2 Synthesis of thymidine-derived β-amino acid monomers

Synthesis of a thymidine amino acid monomer protected with Fmoc at the 3′-amino group (**3**) has been reported previously by this group[156] (Figure 2.2). This amino acid has also been synthesised by other groups; however, each reported synthesis varies in both the chemical route to the product and the intended use of the monomer.[157,158]

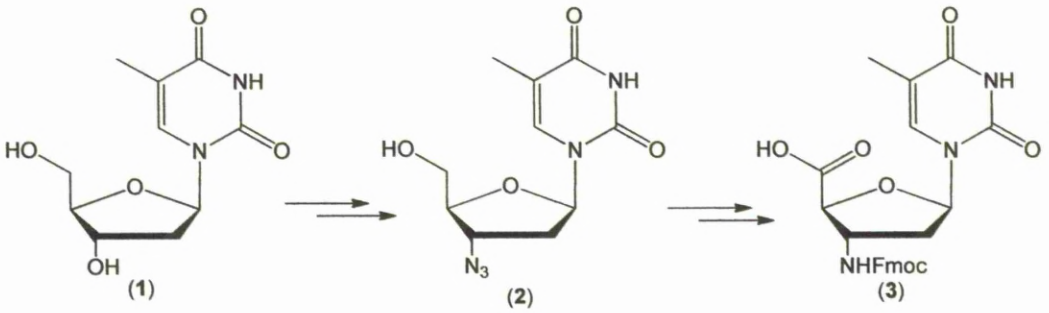


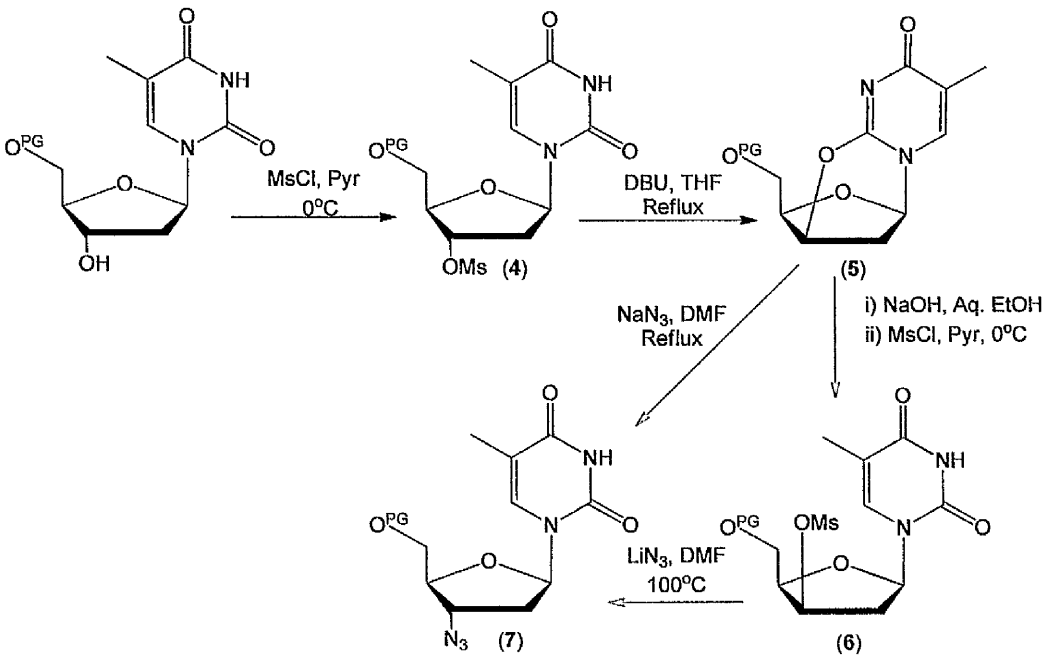
Figure 2.2: Synthesis of the target thymidine β-amino acid monomer (**3**)

In this case, the monomer was intended for use with solid-phase peptide synthesis conditions, but was deemed to also be suitable for coupling reactions in solution. The synthesis of this monomer utilised the pharmaceutically important compound 3′-azido-3′-deoxythymidine (AZT, Figure 2.2, (**2**)). AZT was one of the first approved treatments for HIV and is still used in combination with other drugs today. Consequently, this compound has been studied intensively in order to find efficient routes of synthesis.[159]

Routes to AZT, including synthesis from carbohydrate derivatives,[160] crotonaldehyde[161] and 3,3-diethoxypropionate,[162] exist in the literature but these can be quite complex. By far the most efficient method of synthesis of this compound begins with the parent nucleoside and goes via preparation of an anhydronucleoside intermediate (Scheme 2.1, **5**). This method employs a double S_N2 reaction at the 3′-position of the deoxyribose ring, with overall retention of stereochemistry. The second of these reactions is carried out using an azide nucleophile to introduce this functionality at the 3′-position (**7**).

Anhydronucleosides have been obtained by a number of methods all of which rely on transformation of the 3′-hydroxyl group into a good leaving group (Scheme 2.1). Activation of the 3′-hydroxyl group can be achieved either *in situ,* immediately followed by cyclisation in a 'one-pot' reaction, or the activated intermediate can be isolated and purified before cyclisation is carried out. Generation of the activated

intermediate *in situ* is usually performed using either fluorinating reagents[163] or in a Mitsunobu-style reaction[164] and the most commonly used isolatable intermediates are 3'-*O*-methansulfonyl derivates (**4**).[165] The anhydronucleoside intermediate can then be directly opened by azide[165] or, as demonstrated by Horwitz[166] and Lin,[167] a second 3'-xylo-*O*-methanesulfonyl derivative (**6**) can be synthesised and isolated, prior to displacement with azide.

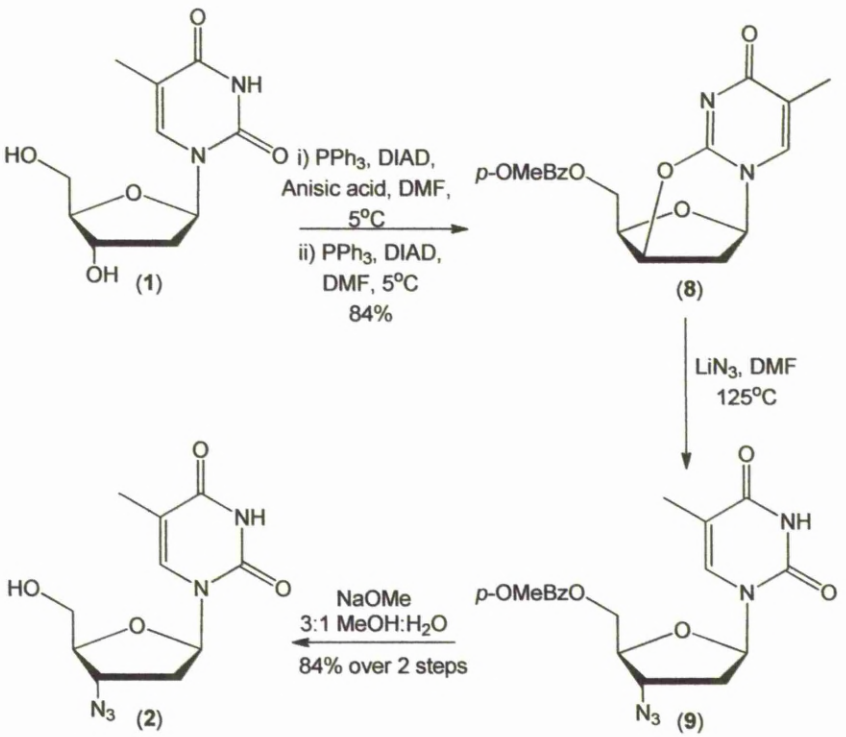


Scheme 2.1: Routes to AZT using direct opening of anhydrothymidine[165] (blue) and isolation of second active intermediate as demonstrated by Horwitz[166]/Lin[167] (pink).

For our investigations, the method of AZT synthesis chosen was the elegant double Mitsunobu reaction developed by Czernecki *et al.* (Scheme 2.2).[164] The initial step in this reaction was protection of the 5'-hydroxyl group using *p*-methoxybenzoic acid and this was followed immediately by a second Mitsunobu reaction to give the 5'-*O*-protected -2, 3'-anhydrothymidine target (**8**). The white powder obtained required no further purification after washing with plenty of diethyl ether and was afforded in high yield (84%). Indications of successful 2,3'-anhydro formation included an unusually high ^{1}H NMR shift for H3' at 5.40 ppm (thymidine H3' appears at 4.10 ppm) and a downfield shift of C3' in the ^{13}C NMR spectrum from 71.0 ppm in thymidine to 77.5 ppm in the 2, 3'-anhydrothymidine compound.

Ring opening of the 2, 3'-anhydronucleoside (**8**) was achieved by treatment with lithium azide in anhydrous DMF at 125°C. Lithium azide was chosen due to its

enhanced solubility in polar organic solvents compared to other azide salts such as sodium azide.[168] Anhydrous lithium azide is not commercially available, but can be easily prepared from a salt exchange reaction between sodium azide and lithium sulfate. The lithium azide and sodium sulfate by-product can be difficult to separate entirely and, for this reason, a large excess of the prepared lithium azide was used in this reaction.



Scheme 2.2: Synthesis of AZT from thymidine (*p*-OMeBz = *para*-methoxybenzoyl)

Azides show a characteristic peak in the IR spectrum at around 2100cm^{-1}. Analysis of the product (**9**) by IR gave a peak at 2103cm^{-1}, which confirmed the presence of an azide group. This was supported by further analysis. The ^{1}H NMR spectrum confirmed an upfield shift of the H3′ proton from 5.40 ppm in the starting material (8) to 4.32 ppm in the azido compound and the ^{13}C NMR spectrum gave an upfield shift for C3′ from 77.5 ppm in the starting material to 61.6 ppm in the product. Comparison of the ^{13}C NMR spectrum of the azido compound to that of the parent nucleoside, thymidine (**1**), also pointed positively to a successful reaction. The C3′ signal was again upfield from its original shift reported for (**1**) and this was as expected due to an increased shielding effect of an azide compared to that of a hydroxyl group.

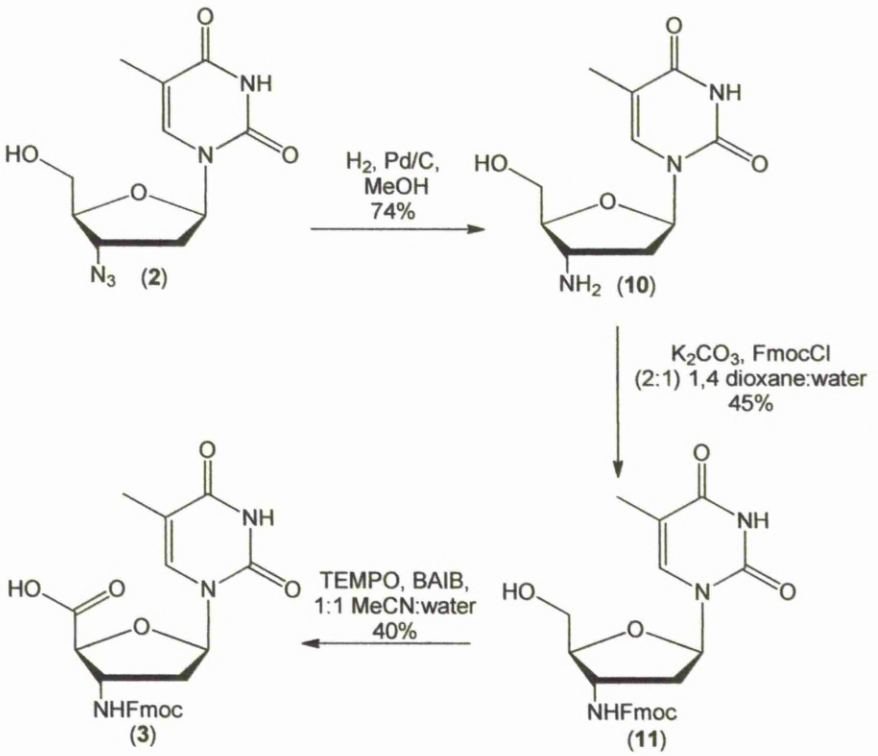
During the reaction, formation of a small amount of a more polar product was observed. This was isolated during purification and was found to be AZT (**2**). This slight loss of the 5′-*O*-protecting group suggested that it was not entirely suited to the harsh conditions required for the opening of the 2, 3′-anhydronucleoside. It was decided tha,t in order to minimise loss of the product, the protected azido nucleoside (**9**) would be carried through the deprotection step crude. This gave an increase in yield from 78% over two steps for the method utilizing isolation of the protected azido intermediate as compared to 84% when the intermediate was not isolated.

Removal of the 5′-*O*-*p*-methoxybenzoyl protecting group of (**9**) was carried out under basic conditions using powdered sodium methoxide in aqueous methanol followed by neutralisation with Amberlite™ H^+ resin. Removal of the protecting group was confirmed by ^{1}H NMR spectroscopy by the disappearance of the characteristic aromatic proton signals associated with the protecting group. This route afforded AZT (**2**) in a 45% yield over 3 steps starting from thymidine (**1**).

The next step in the synthesis of the protected thymidine amino acid was reduction of the azide moiety (Scheme 2.3). Although many reduction methods exist, and some of these will be discussed in more detail later in this chapter, the method chosen here was catalytic hydrogenation using palladium on activated carbon. This provided a simple and efficient route to the 3′-amino nucleoside (**10**) in a 74% yield and without the need for further purification. Analysis of the compound by IR spectroscopy indicated successful reduction due to disappearance of the characteristic azide peak and this was also confirmed in the ^{1}H NMR spectrum of compound (**10**) by an upfield shift of the H3′ proton of around 1 ppm.

Protection of the 3′-amino nucleoside with Fmoc was then carried out by treatment of (**10**) with potassium carbonate in a 33% aqueous 1,4-dioxane solution followed by addition of Fmoc chloride at low temperature. Addition of water to the solution was expected to afford the hydrophobic compound (**11**) as an amorphous solid which could simply be isolated by filteration from the reaction mixture. However, it was found that a sticky solid formed rather than the expected powder so extra purification by flash chromatography was required to obtain the product in sufficient purity for analysis. It was observed, in the ^{1}H NMR spectrum, that the deshielding effect of the protecting group caused a significant downfield shift of the H3′ proton; from 3.33 ppm in the starting material to 4.25 ppm in the product. This, coupled with

the appearance of characteristic aromatic protons, confirmed that protection had been successful.



Scheme 2.3: Route to the final protected thymidine-derived β-amino acid.

The final step in this reaction sequence was the oxidation of the 5'-hydroxyl group of (**11**) to a carboxylic acid. Although nucleosides can tolerate a wide range of oxidation conditions, as will be discussed later in this chapter, mild, metal free oxidations are generally preferred. The milder conditions offer efficient routes to the desired level of oxidation without over reaction or the requirement for complicated purification techniques. For the purposes of this study an oxidation system utilizing the nitroxyl radical, 2,2,6,6-tetramethylpiperidinoxyl (TEMPO), was chosen. TEMPO, along with other nitroxyl radicals, has been identified as a useful reagent for oxidations as it can be used under a variety of conditions and with a number of co-reagents. It is also compatible with a large range of substrates and is highly selective for oxidation of primary alcohols.[169,170] TEMPO can be used in oxidations either in stoichiometric or catalytic quantities, however it has been identified that the active species required for oxidation is not the TEMPO itself but an oxoammonium species.[171] If used stoichiometrically, the oxoammonium species can either be

generated *in situ* using another oxidant[172] or can be prepared separately and isolated.[173]

The catalytic use of TEMPO has been intensively studied as the catalytic cycle for regeneration of the radical species has proved to be mechanistically interesting. The mechanism by which the active oxoammonium species is generated depends upon the pH of the reaction. Two pathways for the generation of the oxoammonium intermediate and regeneration to the radical species exist; below pH 3 disproportionation is favoured and above pH 3 *syn*-proportionation is the favoured cycle (Figure 2.3).[169] Disproportionation is induced by acidic conditions where two molecules of the radical species generate one molecule of the active oxoammonium salt and one molecule of a hydroxylamine by-product. Once the oxoammonium salt has oxidised the alcohol, a second molecule of the hydroxylamine by-product is formed. Both are converted back to the radical species by the co-oxidant so that the cycle can begin again until full oxidation of the alcohol has taken place.

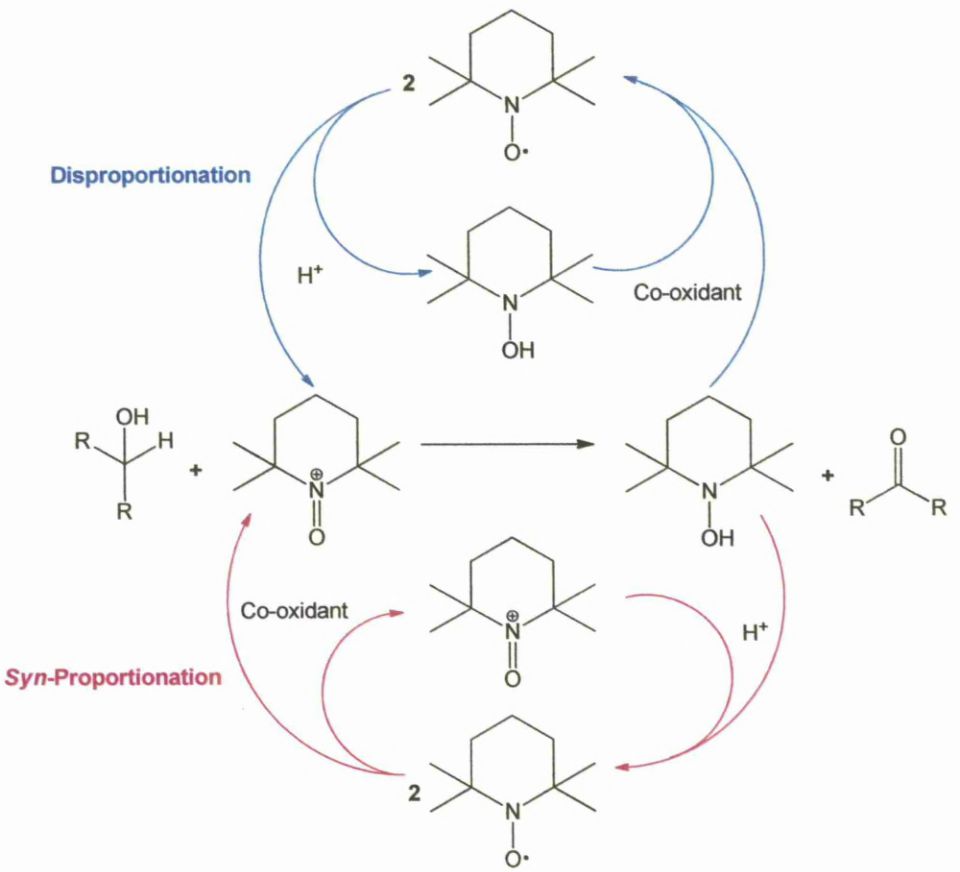


Figure 2.3: TEMPO oxidation catalytic cycles showing disproportionation (blue) and *syn*-proportionation (pink).

Above pH 3, *syn*-proportionation of the oxoammonium and hydroxylamine species occurs to give two molecules of TEMPO radical which can then be oxidised by the co-oxidant to the active oxoammonium salt. The mechanism for the formation of the aldehyde is not yet fully understood. However, three intermediates have been postulated (Figure 2.4).[169]

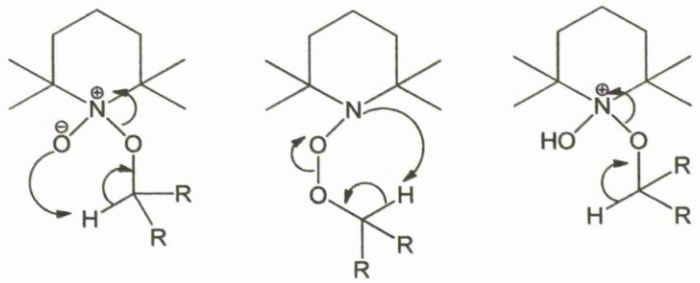
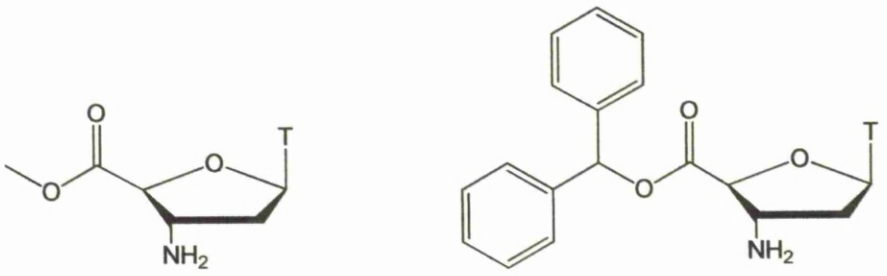


Figure 2.4: Possible intermediates in the TEMPO oxidation.

For this particular conversion, the co-oxidant chosen was *bis*-acetoxyiodobenzene (BAIB) as this can play a dual role of co-oxidant and proton source. Acetic acid is generated during a ligand exchange reaction between BAIB and the substrate and, for this reason, it is thought that this reaction is likely to follow the disproportionation route. Treatment of the protected 3′-amino nucleoside (**11**) with TEMPO/BAIB in 50% aqueous acetonitrile produced the desired 5′-carboxylic acid (**3**) in a 40% yield. The compound precipitated from the reaction liquors upon formation and required no further purification. The disappearance of the two distinctive H5′ proton signals from the ^{1}H NMR spectrum and shift of the C5′ peak from 62.7 ppm to 172.5 ppm in the ^{13}C spectrum allowed for easy characterisation of the compound.

Although a yield of less than 50% of the protected amino acid was obtained, the fate of the other half of the material remains uncertain. Analysis of the reaction liquors showed only the presence of the expected by-products, acetic acid and iodobenzene. It has been identified that hypervalent iodine compounds, such as BAIB, can potentially act as an ozone equivalent and mediate carbon double bond cleavage in the presence of water.[174] No physical evidence was found to support this theory through analysis of the reaction by-products. However, it has been demonstrated that the double bond in the thymidine base is susceptible ozonolysis.[175]

A second thymidine monomer was designed with protection of the 5′-carboxylic acid group and free 3′-amino to allow for coupling in solution (Figure 2.5). Although monomers with 5′-methyl ester protecting groups have been reported in the literature and utilised in coupling reactions performed in solution,[157] removal of these groups and regeneration of the acid can be extremely difficult and require harsh conditions. In addition, the methyl group does little to aid the solubility of these polar compounds. In order to enable synthesis of longer oligomers, a protecting group that would behave in an orthogonal fashion to the Fmoc group was required. Therefore, the benzhydryl group was chosen as removal requires only a simple hydrogenation reaction. It was also anticipated that this protecting group would improve the compound's solubility in much the same way as a trityl group.



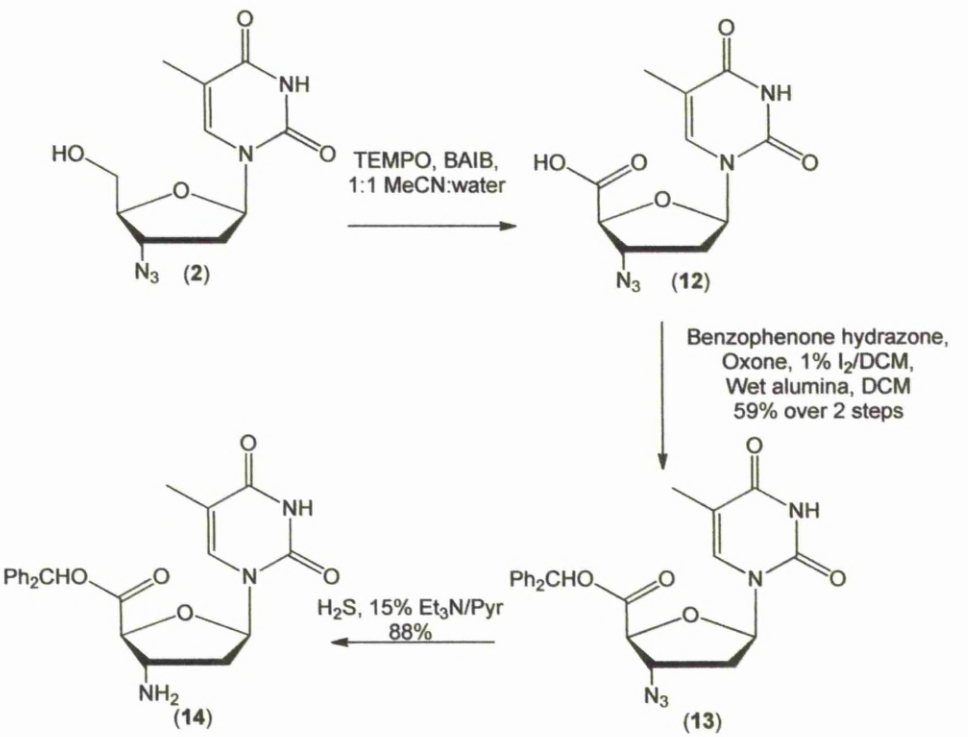
Methylester **Benzhydrylester**

Figure 2.5: 5′-Carboxylic acid protecting groups (T=thymine).

Synthesis of the benzhydryl-protected monomer (Scheme 2.4) began with the previously discussed AZT (**2**) and oxidation of the 5′-hydroxyl group to give thymidine 3′-azido acid (**12**). Oxidation was carried out using the same TEMPO/BAIB procedure described previously. Purification at this stage proved difficult due to the surprising solubility of the compound in most organic solvents and attempts to purify the azido acid (**12**) by trituration with diethyl ether and hexane gave only a sticky solid. As the high polarity of this compound was expected to make chromatographic purification difficult, it was decided to carry the crude azido acid through to the next step involving preparation of the benzhydryl derivative (**13**).

Many methods for the introduction of the benzhydryl protecting group can be found in the literature ranging from reaction of carboxylic acids with diphenylmethyl phosphates,[176] silver carboxylates with diphenylmethyl chloride[177] and oxidation of benzophenone hydrazone to generate the active intermediate

diphenyldiazomethane.[178] The simplest and most efficient route to introduce this protecting group seemed to be using the latter of these methods, however diazo compounds are well known to be explosive and undergo degradation upon storage[179]. It was therefore decided to follow a simple one-pot procedure devised by Curini *et al.* in which the reactive diphenyldiazomethane intermediate was generated *in situ* by oxidation with Oxone® and reacted immediately.[180]



Scheme 2.4: route to 5′-protected thymidine amino acid (**14**).

Thus, a cooled solution of the azido acid (**12**) and benzophenone hydrazone in the presence of catalytic iodine was treated with Oxone® supported on wet alumina.[181] Once the reaction was judged to be complete by tlc, the unwanted solids were filtered off and a crude residue obtained after removal of the reaction solvents. Purification was carried out by simple column chromatography to give benzhydryl ester (**13**) in a 59% yield. The formation of benzhydryl ester (**13**) was confirmed by ^{1}H NMR spectroscopy as the spectrum recorded showed the appearance of aromatic peaks integrating as 11 protons. The singlet peak of the individual hydrogen atom could not be distinguished clearly due to overlap with the aromatic protons.

The final step in the route was reduction of the azide moiety; however, a simple hydrogenation could not be used in this case due to the sensitivity of the protecting group to these conditions. Many other methods exist in the literature for the reduction of azides including reductions with lithium aluminium hydride,[182] $InCl_3$[183] and $Fe/AlCl_3$ or $Fe/BiCl_3$.[184] However, use of heavy metals can lead to problems with purification and the harsh conditions sometimes required are not compatible with nucleosides.

The most attractive methods for the reduction of azides in nucleoside chemistry are those which offer mild conditions with a simple work-up and purification procedure. These include the Staudinger reduction,[185] 1,3-propanedithiol[186] reduction and hydrogen sulfide reduction.[187] Although the Staudinger reduction provides an extremely efficient route to amines from azides, it had been found previously in the group that the iminophosphorane intermediate can be difficult to hydrolyse and removal of the triphenylphosphine oxide by-product can prove problematic. It was therefore decided to reduce the azide using hydrogen sulfide gas (Figure 2.6). The major by-product of this reaction is elemental sulfur which, if required, can be removed by extraction of the desired compound into water; sulfur has extremely poor solubility in aqueous solvents.

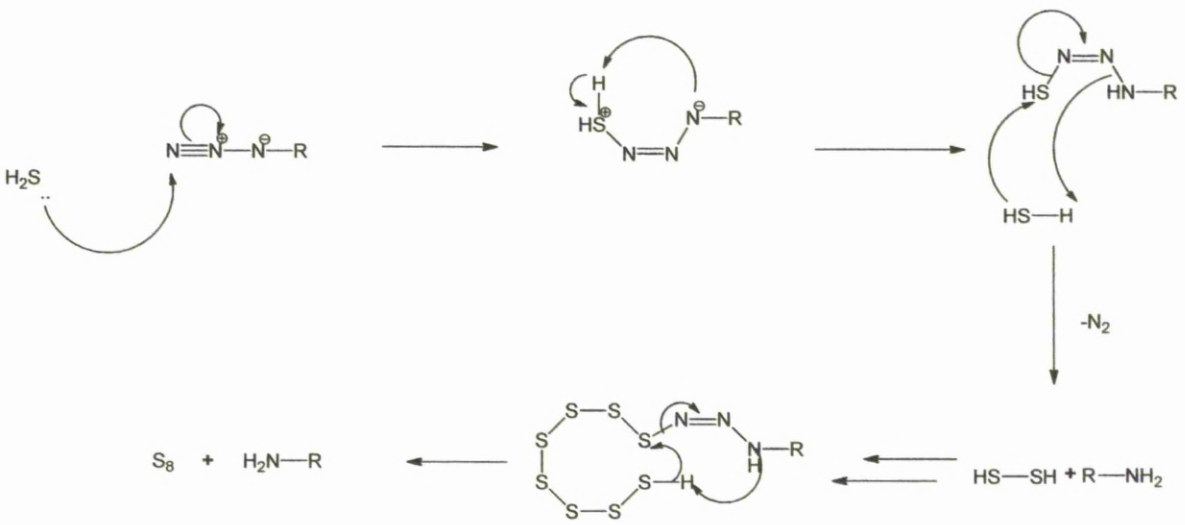


Figure 2.6: Mechanism of azide reduction by H_2S.

Reduction of the protected azido acid (**13**) was performed by bubbling H_2S through a solution of the nucleoside and provided the amino acid monomer (**14**), after purification by flash chromatography in an 88% yield. Generation of the amino

compound was confirmed by 1H NMR through a significant upfield shift of the H3′ peak from 4.36 ppm in the starting material (**13**) to 3.76 ppm in the product (**14**). Use of the monomers (**3**) and (**14**) for coupling in solution will be discussed in Chapter 3.

2.3 Synthesis of a 5-methylcytidine-derived β-amino acid monomer

It was decided to synthesise a monomer derived from cytidine for a number reasons. Firstly, we wanted a monomer that would base pair with guanosine; it was thought that even though there is a methyl group present at the 5-position and so this is not a 'true' cytidine base, this should have no affect on the ability of the base to participate in standard Watson-Crick hydrogen bonding. Secondly, we wanted to examine the feasibility of using AZT as a common intermediate in the synthesis of both pyrimidine nucleoside β-amino acids, thereby reducing the number of steps required to make the cytidine derivative.

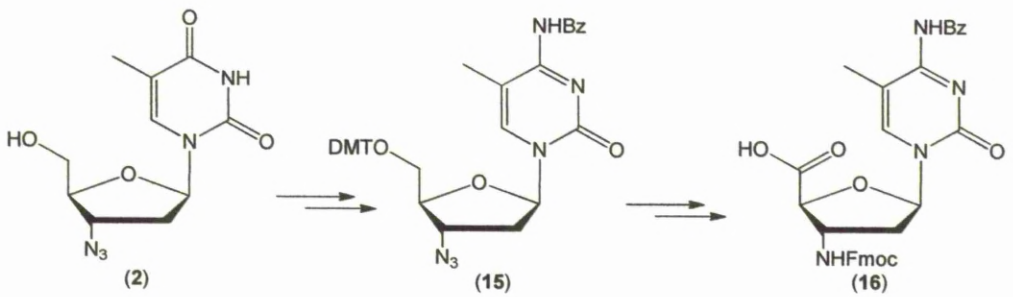
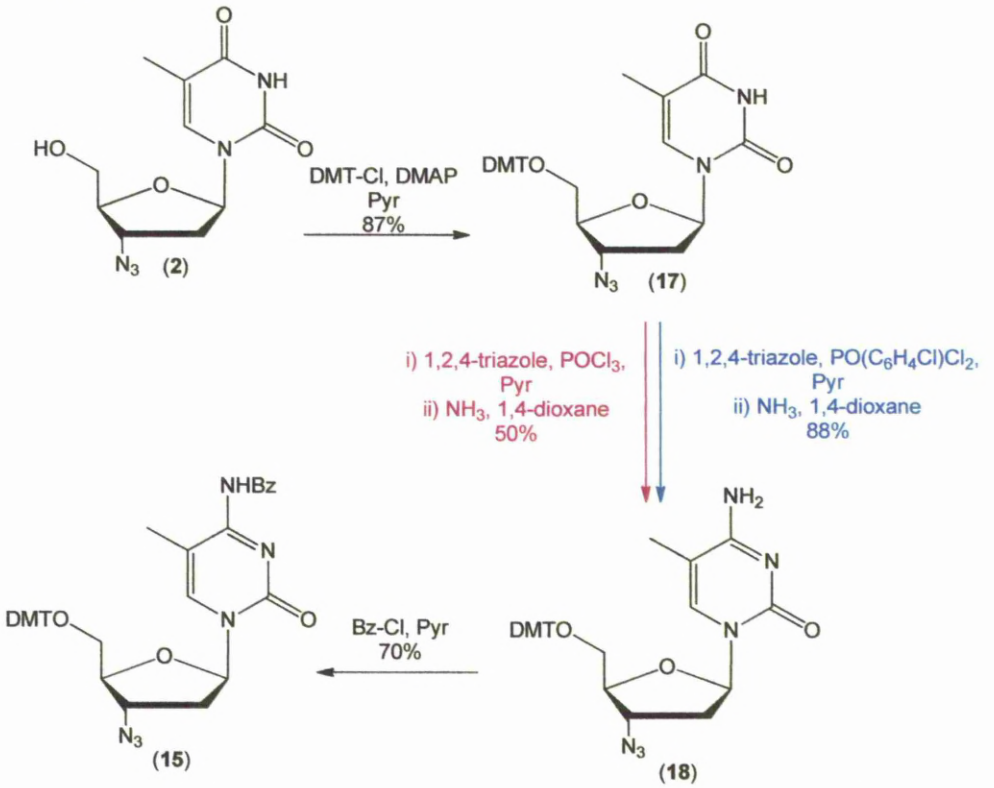


Figure 2.7: Synthesis of the 5-methylcytidine derived β-amino acid monomer (**16**) (DMT= 4, 4′-dimethoxytrityl).

The route to a 5-methyl-2′-deoxycytidine derived β-amino acid monomer (**16**, outlined in Figure 2.7) was designed using a similar strategy to that of the thymidine compound. AZT (**2**) was used as a common intermediate in both synthetic routes and it was thought that a functional group conversion carried out at the 4-carbonyl of the thymidine base would offer an efficient route to a cytosine derived β-amino acid analogue. A more detailed breakdown of the initial steps in this synthesis is shown in Scheme 2.5.

AZT (**2**) was synthesised by the previously discussed route and the 5′-hydroxyl group was protected with a 4, 4′-dimethoxytrityl (DMT) group. The 5′-protected AZT derivative (**17**) was easily accessed through an S_N1 reaction with DMT chloride in

the presence of dimethylaminopyridine (DMAP) in anhydrous pyridine (Scheme 2.5). This afforded the desired compound (**17**) in an 87% yield after purification. The identity of (**17**) was confirmed by ^{1}H NMR by the appearance of characteristic aromatic peaks and a singlet at 3.76 ppm corresponding to the 6 protons of the methoxy groups.



Scheme 2.5: Route to fully protected 3′-azido-2′,3′-dideoxy-5-methylcytidine (**15**).

The first known synthesis of a 5-methylcytidine derivative from thymidine was carried out by Fox *et al.* in 1959.[188] In this procedure, thiation of the 4-position, using phosphorus pentasulfide, and subsequent displacement with ammonia afforded the nucleoside in good yield. However, these conditions are considered to be quite harsh and since this publication many other milder methods for carrying out the same conversion have been developed. These include sulfonylation-amination,[189] chlorination-amination[190] and triazolation-amination.[191] From a survey of the literature, the latter seemed to be the most commonly used method for the conversion and the most efficient. Based on this evidence, it was decided to follow this route.

Initial attempts at the triazolation-amination route proved to give disappointing results. For example, a synthesis reported by Gogoi *et al.*, where triazolation was

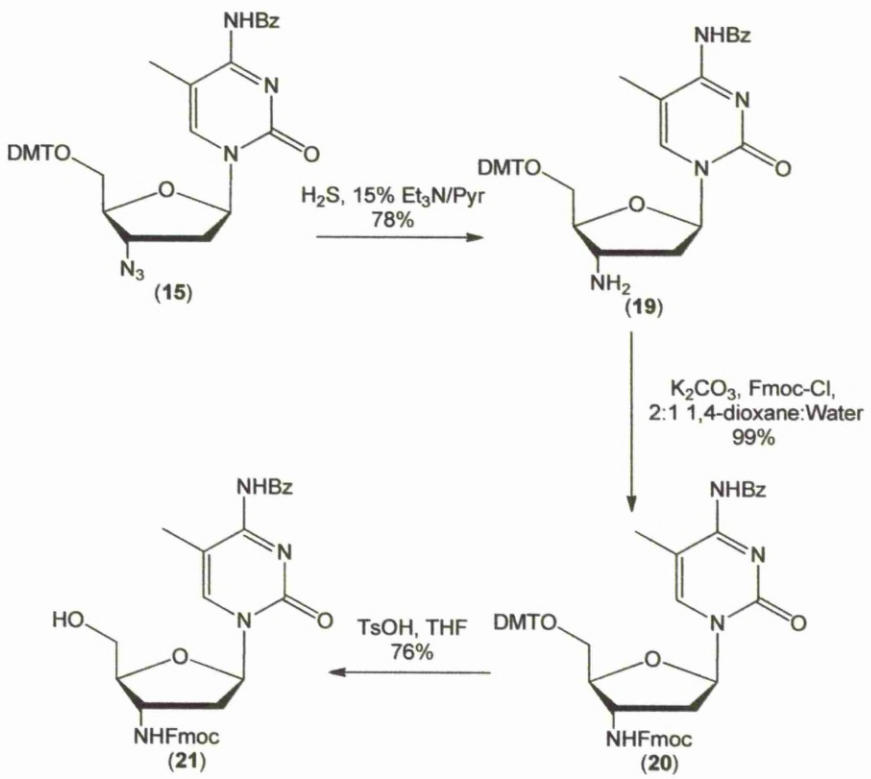
carried out by use of phosphoryl trichloride and 1,2,4-triazole (Scheme 2.5, pink), followed by amination gave a maximum yield of 50%[187] and the reported yield of 96% was not achieved even after multiple attempts. An alternative method utilizing *p*-chlorophenyl phosphodichloridate (Scheme 2.5, blue), published in 1981 by Sung had reported slightly lower yields, but was found to be much more reliable in practice.[192] Thus, the 5′-protected AZT intermediate (**17**) was stirred with 1,2,4-triazole and *p*-chlorophenyl phosphodichloridate in pyridine for three days to give the 4-triazole intermediate. This was immediately treated with an ammonia/1,4-dioxane solution to give the 3′-azido-2′,3′-dideoxy-5-methylcytidine derivative (**18**). After purification, a yield of 88% was obtained which falls in line with the reported yield of 85%. The conversion was confirmed by the appearance of a broad singlet in the ^{1}H NMR spectrum at 8.21 ppm corresponding to the two protons of the NH_2 group and loss of one mass unit in the high resolution mass spectrum.

Protection of the newly formed exocyclic amine of (**18**) required a group that would behave orthogonal to the Fmoc group for the protection of the sugar 3′-amino group later in the synthesis. The *N*-benzoyl group is well established in both PNA[193] and oligonucleotide[194] synthesis and was deemed suitable to withstand any of the conditions that would subsequently be used in this synthesis.Thus, 3′-azido-2′, 3′-dideoxy-5-methylcytidine (**18**) was treated with benzoyl chloride in pyridine and after simple work-up and purification the *N*-4-benzoylated compound (**15**) was obtained in 70% yield. Although the appearance of more aromatic peaks was observed in both the^{1}H and ^{13}C NMR spectrum it was difficult to ascertain from this alone if the correct compound had been formed. However, an accurate high resolution mass spectrum was obtained which confirmed that the benzoyl group was indeed present.

From this point the synthetic route was planned to proceed in a similar manner to that which had been developed for the preparation of the thymidine amino acid (**3**) (Scheme 2.6). Therefore compound (**15**), with protection on both the exocyclic amine and 5′-hydroxyl group, was subjected to reduction by simple hydrogenation. However, this reaction failed to give any of the amine product expected. Thus, it was decided to employ the previously used H_2S method of reduction which gave the 3′-amino nucleoside (**19**) in a 78% yield, after purification by flash chromatography. The identity of the product was confirmed by the appearance of a broad singlet in the ^{1}H NMR spectrum at 8.29 ppm corresponding to 2 protons. However, all attempts to remove the DMT protecting group at this point resulted only in degradation of the product presumably by cleavage of the glycosidic bond. It was,

therefore, decided to continue the synthesis with the 5′-protecting group intact as shown in scheme 2.6.

Protection of the newly formed amine with an fmoc group was, therefore, carried out in the same manner as discussed previously for the thymidine amino acid (**10**). The fully protected nucleoside (**20**) was afforded in near quantitative yield. Successful protection was confirmed by the peak shift for H3′ in the ^{1}H NMR spectrum, from 3.77 ppm in the starting material,I to 3.93 ppm in the product and the appearance of the characteristic Fmoc peaks at 4.50 ppm and 4.98 ppm. Treatment of compound (**20**) with tosic acid in THF, to remove the 5′-DMT protection, proceeded smoothly and the 5′-deprotected nucleoside (**21**) was afforded in good yield ready for the final oxidation step. The product was confirmed by loss of 13 protons from the aromatic region of the ^{1}H NMR.



Scheme 2.6: Synthesis of 3′-*N*-protected amino nucleoside.

Unfortunately, at this stage, attempts to generate the final protected amino acid (**16**) using the preferred method of TEMPO oxidation failed (Figure 2.8). The presence of

the hydrophobic Fmoc group rendered compound (**21**) completely insoluble in aqueous acetonitrile, the optimal solvent system required for this oxidation.

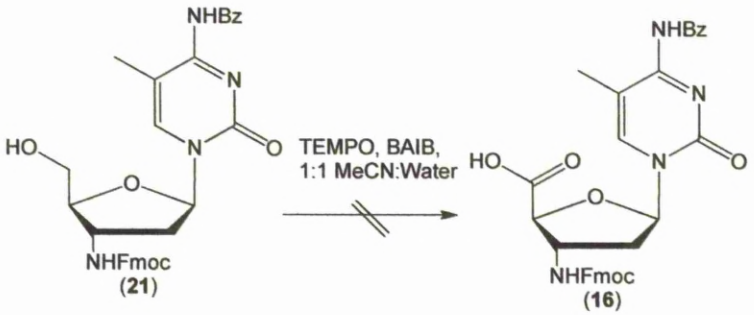


Figure 2.8: Oxidation.

This unexpected set back prompted an investigation into other methods of oxidation. Although mild metal free methods of oxidation are preferred within the area of nucleoside chemistry, some traditional transition metal based procedures remain an option. Oxidations utilising $KMnO_4$ are not usually considered suitable for nucleoside transformations due to the harsh conditions required,[195] but pyridinium dichromate has been shown to be useful in the oxidation of certain nucleosides.[196] More popular methods of oxidation incorporating the use of heavy transition metals include $RuCl_3$ with persulfate[197] and O_2 insertion over Adams catalyst (PtO_2) in aqueous $NaHCO_3$.[198] However, transition metal based procedures have a number of potential problems associated with them. These include problems with the co-ordination of metal catalysts to the nucleobase, making complete removal of the catalyst difficult to achieve. Despite these potential issues, the $RuCl_3$ and PtO_2 methods remain widely used in nucleoside chemistry. However, when attempts were made to oxidise the 3′-protected amino nucleoside (**21**) using $RuCl_3$ and potassium persulfate no oxidation was observed.

Due to the problems associated with metal based oxidations, organic, metal-free methods have increased in popularity. A disadvantage of a lot of these oxidations is that they are often not selective for the oxidation of primary alcohols over secondary and the product formed is usually an aldehyde so a further oxidation is required to generate the carboxylic acid. A popular method of oxidation in organic chemistry is the Swern oxidation involving the use of activated DMSO in combination with an electrophiles.[199] Common electrophiles used include oxalyl chloride,[200] acetic anhydride[201] and thionyl chloride.[202] Moffat *et al.* have shown that this type of reaction using DMSO in combination with DCC is a good way of generating

aldehydes in nucleoside chemistry[203,204] and this functionality can then be further oxidised to the carboxylic acid.

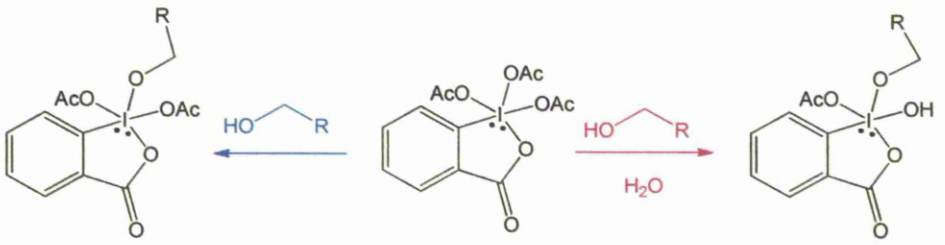


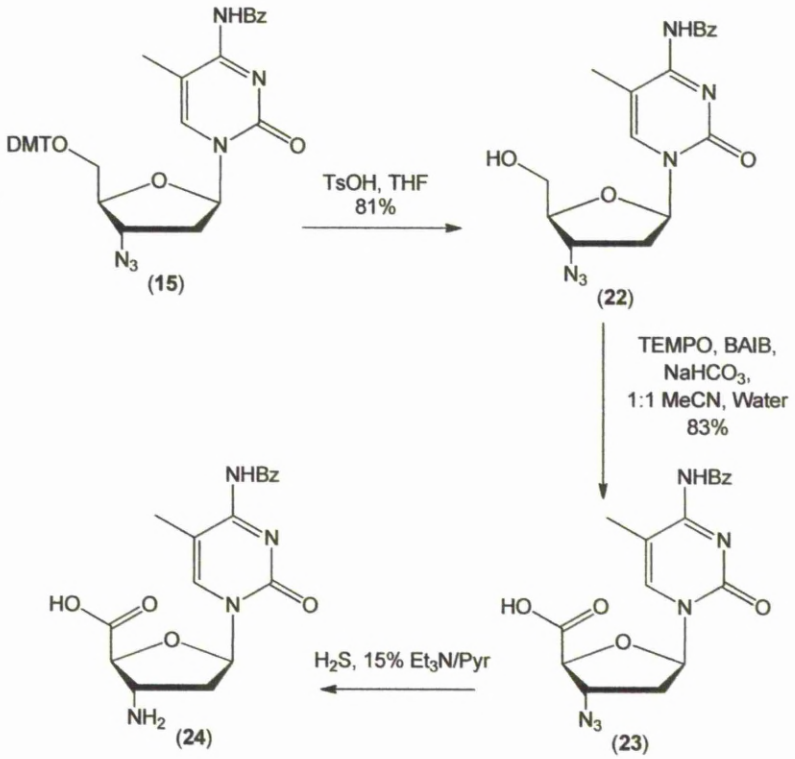
Figure 2.9: Intermediates formed during the Dess-Martin oxidation with water (pink) and without water (blue).

Another popular method for carrying out metal free oxidations is the Dess-Martin procedure. This method employs a cyclic periodane and offers an efficient route to an aldehyde with short reaction times and simple purification. The Dess-Martin periodane is highly selective for the oxidation of alcohols over other oxidisable groups, for example furan rings and sulphides. However, it is not selective for primary alcohols over secondary.[205] The mechanism of the oxidation proceeds firstly *via* a ligand exchange at the iodine centre (Figure 2.9) followed by a proton exchange between the substrate and one acetate ligand and elimination of the newly oxidised product. It has been found that the addition of water to the reaction can accelerate the rate of oxidation as substitution of a second acetate ligand by a hydroxyl group takes place. It is thought the electron-donating ability of the hydroxyl group causes a lengthening of the remaining iodine-acetate bond (Figure 2.9). As a consequence of this the rate of proton exchange between the acetate ligand and substrate is increased and hence the oxidised substrate is released more rapidly.[206]

Attempts to oxidise using standard conditions of the Moffat and Dess-Martin procedures successfully produced the aldehyde product. However, no clear oxidation to the acid could be achieved. It was therefore decided to return to an earlier stage of the synthesis, namely the DMT protected azide (**15**), and investigate an alternative route (Scheme 2.7).

Removal of the DMT protecting group from the azido derivative (**15**) proved difficult, as there was a narrow window between complete removal and product degradation, presumably through cleavage of the glycosidic bond. When the compound was subjected to treatment with tosic acid, reaction times of longer than 10 minutes resulted in a significant reduction in yield and therefore the reaction required careful

monitoring. Fast efficient workup and rapid purification were also required to obtain the best yields. The 5′-deprotected nucleoside (**22**) was obtained in 81% yield and its formation was confirmed by ^{1}H NMR through the loss of 13 aromatic proton signals.

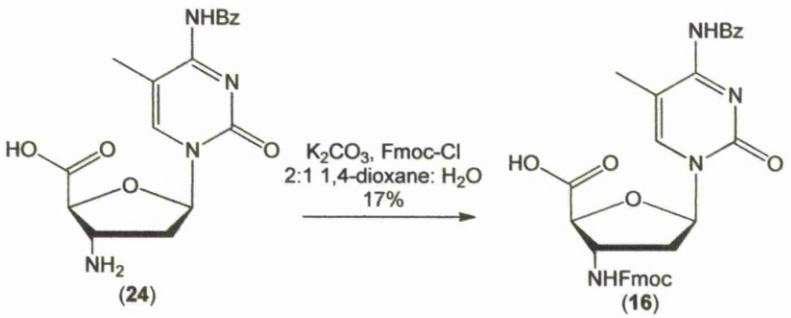


Scheme 2.7: Alternative synthesis of 2′, 3′-dideoxy-5-methylcytidine β-amino acid (**24**).

Oxidation of the 5′-position of nucleoside (**22**) was now successfully carried out using TEMPO. However, due to the apparent sensitivity of the compound to the presence of acid, it was decided to carry out the reaction under the conditions required for *syn*-proportionation. This was achieved by addition of $NaHCO_3$ to bring the reaction mixture to a higher pH. The oxidised product (**23**) was obtained as a brown powder in 83% yield, after trituration from a mixture of acetone and diethyl ether. Successful oxidation was confirmed by a loss of the two signals in the region of 3.76-4.02 ppm in the ^{1}H NMR of the starting material (**22**) corresponding to the two H5′ protons.

Insolubility of azido acid (**23**) in many solvents, including those favoured for hydrogenation, meant that simple palladium-catalysed reduction of the azide moiety

was a less attractive option here. Therefore, reduction was carried out using H_2S as previously described. The high polarity of the amino acid was anticipated to make purification by flash chromatography very difficult so the crude product (**24**) was instead partially purified by dissolution in water to precipitate the elemental sulfur. The identity of the amino acid (**24**) was confirmed by high resolution mass spectrometry.



Scheme 2.8: Protection of 3'-amino group.

Initial attempts to protect the 3'-amino group of (**24**) with an Fmoc protecting group under the same conditions employed with the thymidine amino acid (**10**) gave only a diprotected Fmoc derivative which was confirmed by both NMR and mass spectrometry. This was a surprising result as no di-protection had been observed in the synthesis of the analogous thymidine compound. It appeared that the second Fmoc group had been incorporated at the *N*-4-position, although there was no physical evidence to support this theory. The reaction was therefore attempted with only 1 equivalent of Fmoc-chloride at reduced temperature. Although this change in the conditions slowed down the rate of reaction, significant formation of di-protected product was again observed and it was decided to stop the reaction before completion. The mixture of starting material, mono-protected and di-protected products was separated by flash chromatography and a yield of 17% of the mono-protected target compound (**16**) was obtained (Scheme 2.8). The product was confirmed by the appearance of characteristic Fmoc peaks at 4.53 ppm and 4.70 ppm in the ^{1}H NMR spectrum.

Although the desired protected β-amino acid (**16**) had been successfully synthesised, the low yields obtained from the final protection step suggests that methods for the introduction of the Fmoc group into these β-amino acids requires further study. In addition, related work performed in parallel within the group suggested that the benzoyl group may not be entirely suited to the synthesis and

deprotection of the β-peptides. This is discussed in more detail with regard to the synthesis of the adenine derived β-amino acids. The use of this 2',3'-dideoxy-5-methylcytidine β-amino acid monomer (**16**) for assembly of oligomers in solution will be discussed in Chapter 3.

2.4 Synthesis of an adenosine-derived β-amino acid monomer

2.4.1 Proposed route and strategy for base protection

It was clear that a logical extension of this project would be to synthesise β-amino acids of the purine nucleosides. 2'-Deoxyadenosine was chosen as the next target as it is well known in nucleoside chemistry that guanosine derivatives can be notoriously difficult to manipulate. It was hoped that designing an efficient route to the 2'-deoxyadenosine β-amino acid would provide us with a basis for a route to the 2'-deoxyguanosine β-amino acid. It was also hoped that, in the future, interactions of homo-oligomers of the adenosine and previously prepared, thymidine-derived β-amino acids could be examined.

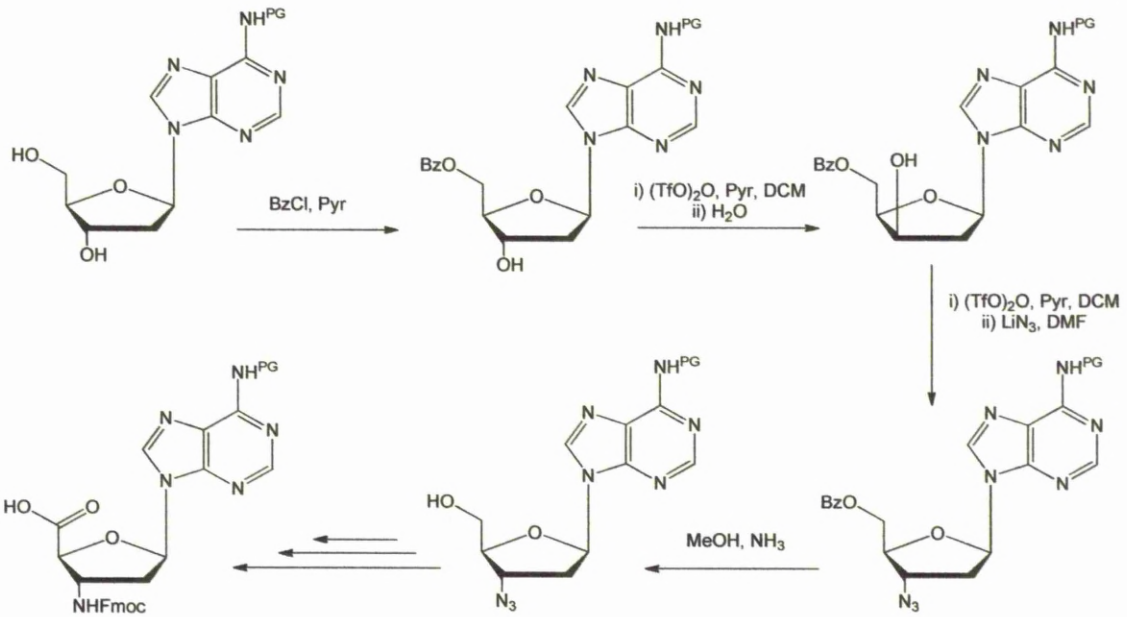
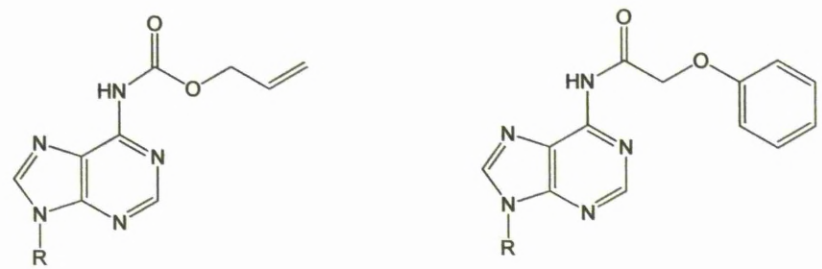


Figure 2.10: Proposed route to an adenosine derived β-amino acid (PG= protecting group, Bz= Benzoyl).

The major difference between the synthesis of the purine and pyrimidine β-amino acids is that, in the case of the purines, an anhydronucleoside cannot be utilised as

an intermediate for the introduction of the azido group. This makes the synthesis of the purine monomers longer and more complex than the pyrimidine monomers. Based on the knowledge gained during the synthesis of the pyrimidine monomers and research into other nucleoside derived structures, an initial route to the adenosine-derived β-amino acid was designed (Figure 2.10). In the chosen, route a strategy employed by Herdewijn for the inversion of stereochemistry at C3′ was deemed to be the most suitable for this synthesis.[207] The initial inversion of stereochemistry requires activation of the 3′-hydroxyl group followed by displacement induced by neighbouring group participation from a 5′-benzoyl protecting group. This gives an unusual bicyclic intermediate which upon treatment with water is opened up to give inversion of configuration at C3′.

The successful implementation of the reaction scheme shown in figure 2.10 depended on the choice of a suitable protecting group for the exocyclic amino group of adenine. It had to be efficiently introduced and stable to the conditions required for subsequent removal of the 5′-benzoyl group. Previous work in the group had shown that benzoyl protection of the base is not entirely suited to the synthesis of the nucleoside β-amino acid oligomers as it can be difficult to remove from the final products without some degradation of the peptide bonds. It was therefore decided to examine two groups which could be removed under milder conditions, the allyloxycarbonyl (Alloc) and phenoxyacetyl (PAC) group (Figure 2.11).



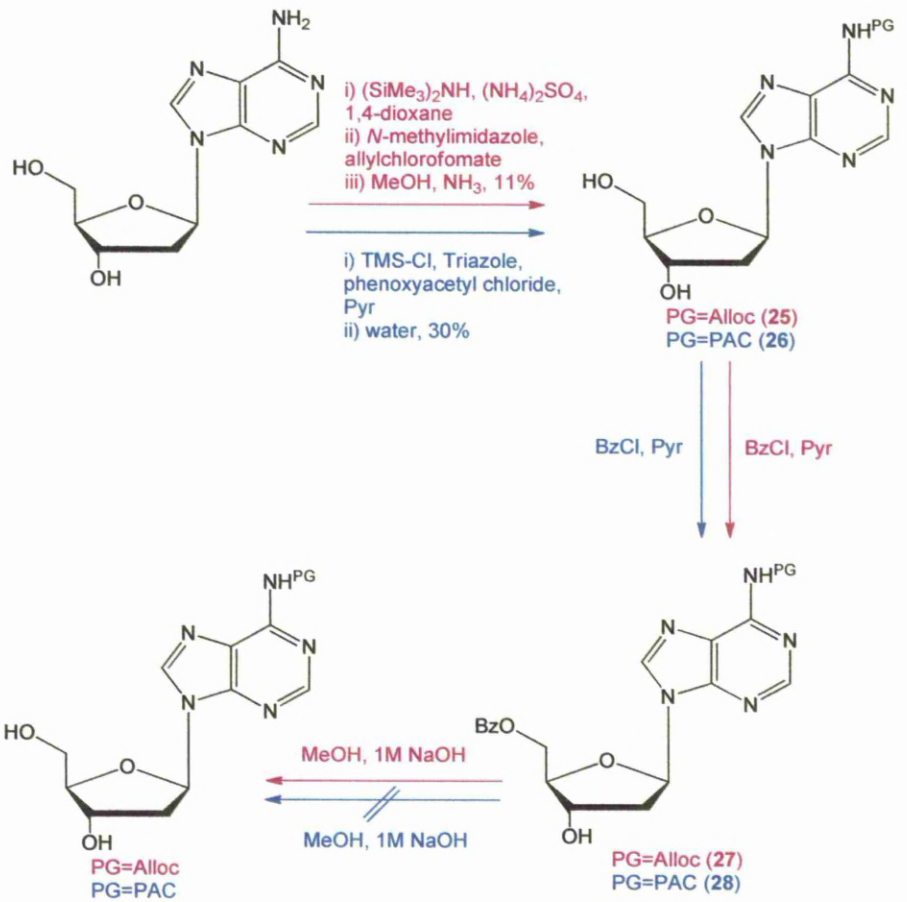
Allyloxycarbonyl (Alloc) **Phenoxyacetyl (PAC)**

Figure 2.11: Protecting groups.

In order to determine which of these protecting groups would be best suited to the route devised, it was decided to synthesise both 5′-benzoyl-*N*-6-protected nucleosides (**27** and **28**).

Protection of the N^6-position of the adenine base was initially investigated using the Alloc group following the literature procedure of Hayakawa *et al.* (pink, Scheme

2.9).[208] This procedure gave a very poor yield (11%) of the N^6-protected nucleoside (**25**). Importantly, though, it was shown that the Alloc protecting group was stable to the preferred conditions (MeOH/1M (aq) NaOH) required for the subsequent removal of the 5'-benzoyl protecting group. In comparison, the PAC group (blue, Scheme 2.9) was introduced in a slightly higher yield (30%). However, removal of the 5'-benzoyl group with MeOH/ 1M (aq) NaOH resulted in significant loss of the PAC group from the base. With the Alloc group proving to be more suitable for the synthetic route, a more efficient method for the introduction of the Alloc group was required since the route devised by Hayakawa *et al.*gave repeatedly poor yields in our hands.

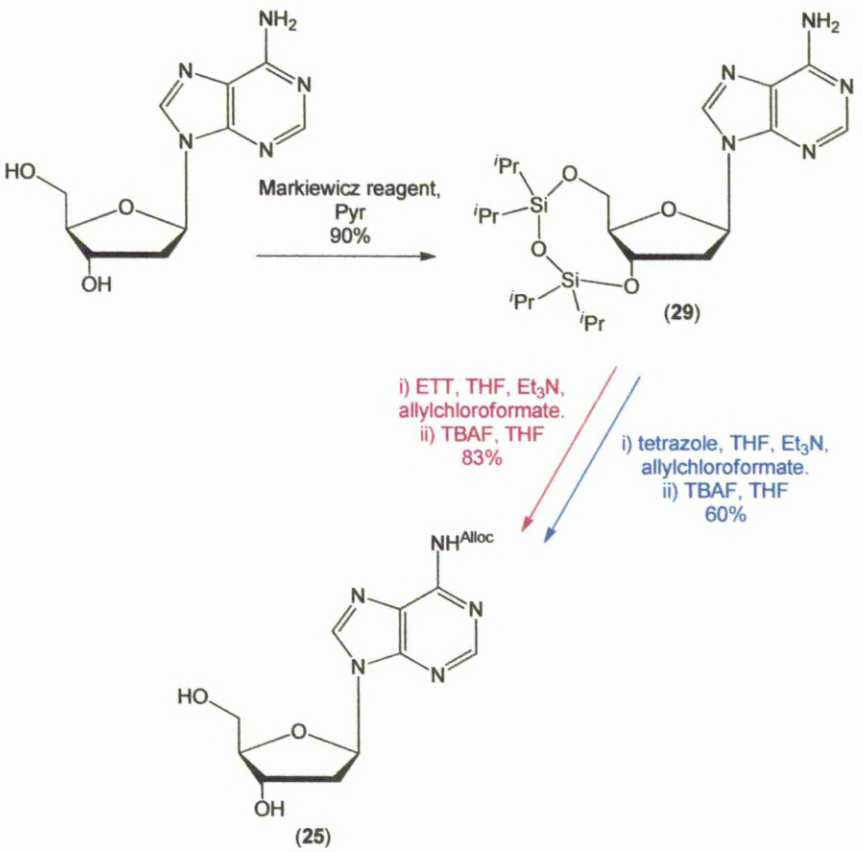


Scheme 2.9: Evaluation of the Alloc and PAC groups for protection of the exocyclic amino group of adenine (Alloc= allyloxycarbonyl, PAC= phenoxyacetyl).

A common reagent used for transient protection of both the 3'- and 5'-hydroxyl groups in nucleoside chemistry is the Markiewicz (TIPS) reagent, 1,3-dichloro-1,1,3,3-tetraisopropyl disiloxane.[209] This reagent is commercially available or can be

prepared in a two step synthesis from trichlorosilane.[210] Protection of the 2'-deoxyribose moiety with this group provides a stable intermediate, unlike the TMS protected derivative, which can be isolated and stored for future use. 2'-Deoxyadenosine was thus treated with the Markiewicz reagent[211] to provide the 3', 5'-TIPS protected nucleoside (**29**) in 90% yield after purification. Protection was confirmed by the presence of a series of peaks between 1.11 and 1.30 ppm in the ^{1}H NMR spectrum corresponding to 28 protons. These peaks were attributed to the alkyl protons directly attached to a silicon atom.

The next challenge was to protect the exocyclic amino group of (**29**). Initially, this was carried out by treating the TIPS protected nucleoside (**29**) with a tetrazole-activated allyloxycarbonyl intermediate which had been prepared separately according to a procedure by Hayakawa.[212] Subsequent removal of the silyl protecting group with TBAF afforded the product (**25**) in a 60% yield; however, multiple purifications by flash chromatography were required to obtain the product in sufficient purity for analysis.



Scheme 2.10: Synthesis of *N*-6-allyoxycarbonyl-2'-deoxyadenosine (**25**).

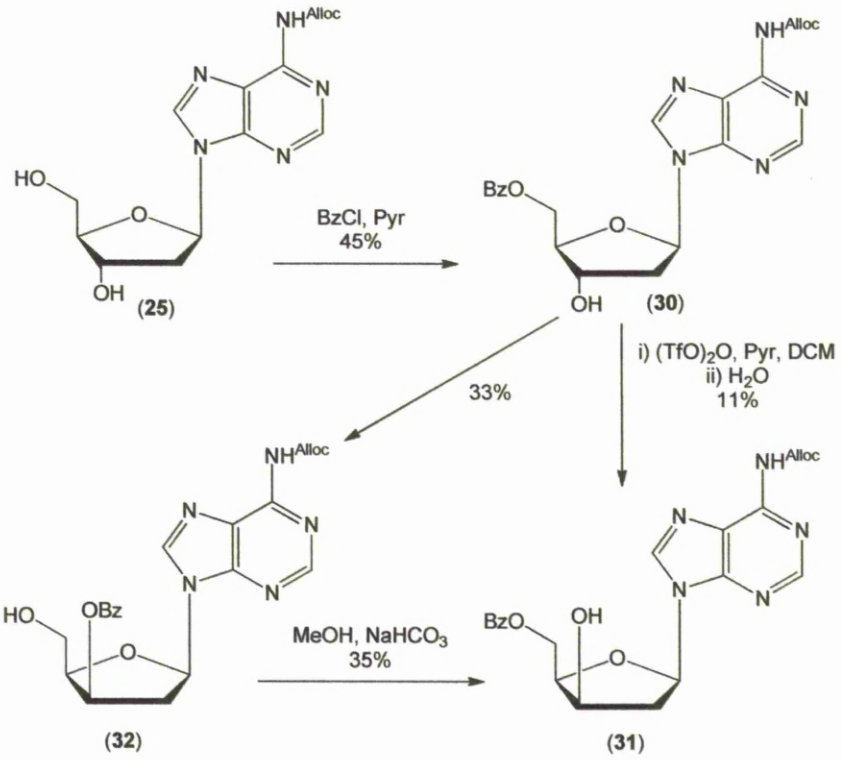
Due to safety issues with regards to the use of tetrazole in large scale syntheses, it was decided to investigate a safer alternative which is commonly used as an activator in the synthesis of oligonucleotides, 5-(ethylthio)-1*H*-tetrazole (ETT). Substitution of tetrazole with this safer alternative gave (**25**) in an improved yield of 83%.In addition,the purification procedure was simplified with the product only requiring one flash column in order to be obtained in sufficient purity for analysis. However, due to reduced reactivity, longer reaction times were required. In both syntheses, the formation of product (**25**) was confirmed by the appearance of characteristic peaks in the ^{1}H NMR spectrum corresponding to the alkene moeity of the allyloxycarbonyl group at 5.29 ppm and 5.39 ppm. These peaks were attributed to the terminal CH_2 protons. The peak at 5.99 ppm was assigned to the alkenyl CH proton.

N-6-Allyloxycarbonyl-2'-deoxyadenosine (**25**) could now safely be produced in large quantities ready for use in further synthesis.

2.4.2 Synthesis of adenosine-derived β-amino acid: Route A

As described previously (Figure 2.10), the synthetic route to the adenosine-derived β-amino acid was to proceed firstly with benzoylation of the 5'-hydroxyl group of (**25**), followed by an inversion of the stereochemistry at C3' and finally displacement of this inverted hydroxyl group with azide to give the required 3'-azido compound with overall retention of configuration.

Selective benzoylation at the 5'-position by treatment of the N^6-protected nucleoside (**25**) with 1 equivalent of benzoyl chloride in pyridine, proceeded to give the desired compound (**30**) in 45% yield (Scheme 2.11). Although a small amount of benzoylation at the 3'-position and dibenzoylation was observed, these side products did not account for the fate of the other 55% of the starting material and this still remains uncertain. Benzoylation was confirmed the appearance of characteristic aromatic peaks in the ^{1}H NMR spectrum of (**30**) at 7.50 ppm, 7.64 ppm and 7.90 ppm corresponding to the *meta*, *para* and *ortho* protons of the benzoyl group, respectively. When the ^{1}H NMR spectrum was recorded in deuterated DMSO, a doublet corresponding to the 3'-O<u>H</u> could also clearly be seen and cross peaks with H3' were observed in a COSY spectrum of the compound. These both indicated that protection had indeed occurred at the 5'-position rather than the 3'-position.



Scheme 2.11: Inversion of stereochemistry at C3′.

The first inversion of stereochemistry was carried out according to Herdewijn's route to 3′-azido-2′, 3′-dideoxyadenosine.[207] The 5′-benzoylated compound (**30**) in a 5% solution of pyridine in DCM was cooled to -30°C and treated with a 10% triflic anhydride/DCM solution, followed by addition of water. An 11% yield of the desired 5′-benzoyl-3′-*xylo*-compound (**31**) was recovered after purification by flash chromatography. However, a 33% yield of the 3′-benzoyl-3′-*xylo* by-product (**32**) was also obtained. These two compounds could be distinguished easily by ^{1}H NMR methods; the 5′-benzoyl-3′-*xylo* nucleoside (**31**) showed only a slight change in shift for the peak assigned to H3′ relative to its position in the spectrum of the starting material i.e., from 4.67 ppm to 4.34 ppm. In contrast, the spectrum recorded of the 3′-benzoyl-3′-*xylo* nucleoside (**32**) exhibited a large downfield shift for this peak from 4.67 ppm to 5.72 ppm.

The formation of the 3′-benzoyl-3′-*xylo* by-product (**32**) is a direct result of the reaction mechanism. Neighbouring group participation by the 5′-benzoyl group in the starting material (**30**) displaces the activated 3′-hydroxyl group to form an unusual bicyclic intermediate (**33**) (Figure 2.12). This intermediate can then be opened in two ways, with the 3′-*O*-benzoate being the favoured product. This is due

to the fact that the primary 5′-hydroxyl group is a better leaving group. A solution to this problem is to isomerise the undesired compound (**32**) back to the 5′-benzoate (**31**) using a weak base such as sodium hydrogen carbonate. This can either be carried out *in situ* or after isolation of the 3′-benzoyl-3′-*xylo* compound. The presence of a weak base catalyses the reformation of intermediate (**33**) and means that the less acidic 3′-hydroxyl group is protonated preferentially. The driving force for the reaction is the improved steric stability of the 5′-benzoate over the 3′-benzoate. Isomerisation afforded the 5′-benzoyl-3′-*xylo* nucleoside (**31**) in a 35% yield.

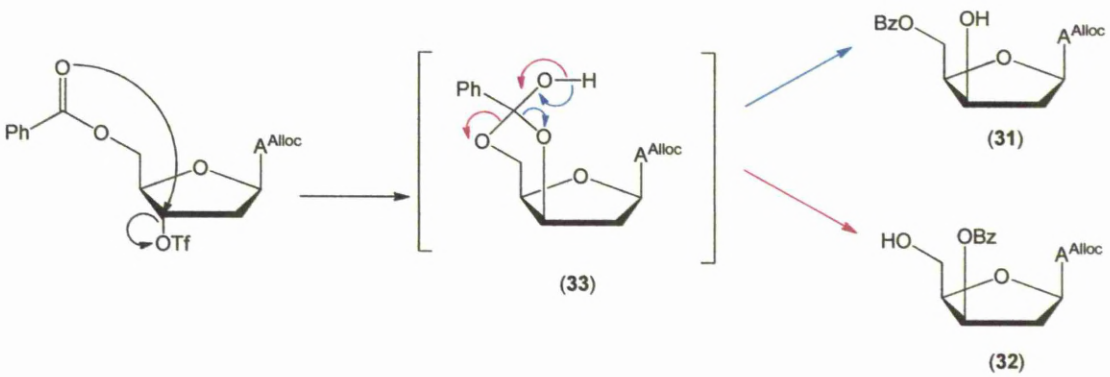


Figure 2.12: Mechanism of xylo nucleoside formation.

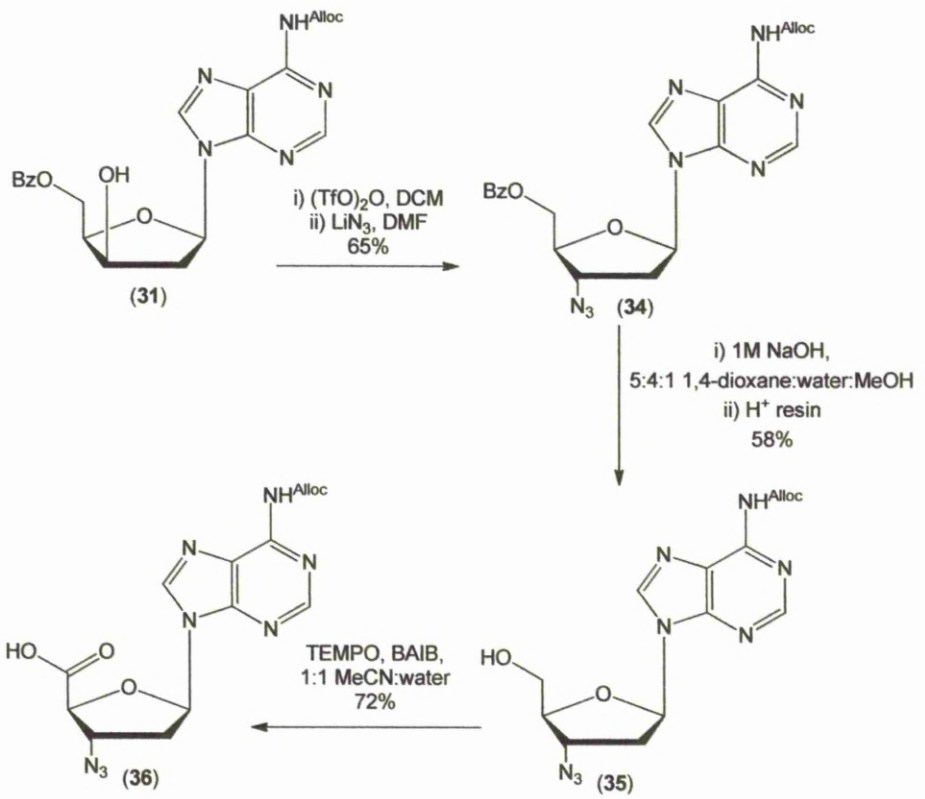
It was apparent from the poor yields obtained for this inversion that this may not have been the ideal method for the transformation and other methods will be discussed later on in this chapter.

The next step in the synthesis was the introduction of the azide moiety. This was carried out by activation of the 3′-*xylo*-OH with triflate followed immediately by displacement with azide in a direct S_N2 attack. This afforded the 3′-azido-5′-*O*-benzoyl-2′-deoxyadenosine derivative (**34**) in 65% yield after purification by flash chromatography (Scheme 2.12). The conversion was once again confirmed by the presence of a strong azide absorption in the IR spectrum at 2090cm^{-1}.

Simple removal of the 5′-benzoyl group under basic hydrolytic conditions afforded 3′-azido-2′-deoxyadenosine (**35**) in 58% yield, after washing with water, and no further purification was required. A small amount of product degradation was observed by tlc, after treatment with the ion-exchange resin, which suggests that the glycosidic bond of this derivative could be extremely acid sensitive. Removal of the

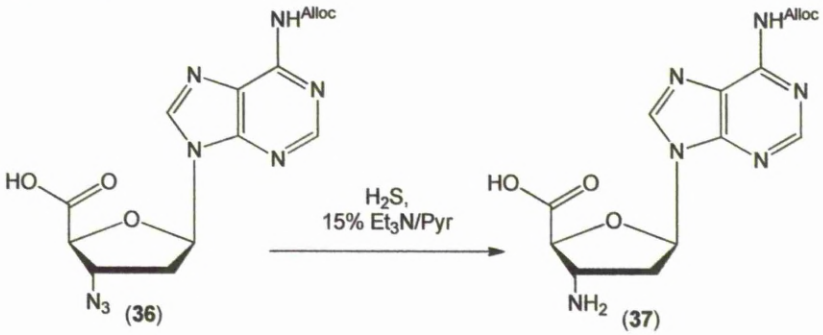
protecting group was confirmed by an upfield shift of the peak assigned to the H5′ protons in the ^{1}H NMR spectrum; from 4.57 ppm and 4.68 ppm in spectrum recorded for the the protected compound (**34**) to 3.82 ppm and 4.03 ppm in the product. There was also a disappearance of the distinctive *ortho*, *meta* and *para* protons of the benzoyl group.

Oxidation of the 5′-position was carried out under standard TEMPO/ BAIB conditions. As with the 5-methylcytidine analogue the adenosine derived azido acid (**36**) did not precipitate from the reaction liquors. However, it was found that the product could be isolated from the reaction by-products by dissolving the crude mixture in methanol and precipitating the acid product through addition of diethyl either. A crop of off-white powder was isolated and found to be the desired product. It was obtained in high purity and 72% yield. Successful oxidation was confirmed by the loss of the signals associated with the two H5′ protons by ^{1}H NMR spectroscopy.



Scheme 2.12: Synthesis of 2′-deoxyadenosine azido acid (**36**).

Reduction of the azide moiety was again carried out by treatment with hydrogen sulfide gas. This reagent was chosen as it was uncertain what effect the presence of palladium would have on the allyloxycarbonyl protecting group which is known to be removed by treatment with tetrakis(triphenylphosphine)palladium in the presence of an allyl scavenger.



Scheme 2.13: Azide reduction.

Due to the polarity of the amino acid (**37**), purification by flash chromatography was not possible. The compound was, therefore, separated from the sulfur by-products by extraction into water and removal of the non-soluble solids by filteration. The formation of compound (**37**) was confirmed by mass spectrometry.

Although multiple attempts at protection of the 3′-amine with Fmoc were carried out, all failed. This combined with the low yields from the inversion step prompted the decision to plan and investigate other routes to the target compound.

2.4.3 Synthesis of adenosine-derived β-amino acid: Route B

A second strategy for inversion of stereochemistry at C3′ involves oxidation *via* a Dess-Martin procedure followed by stereospecific reduction to give the 3′-*xylo*-nucleoside. Stable 2′-deoxy-3′-ketonucleosides were first synthesised and isolated by Robins *et al.* in 1990.[213] Previous efforts to prepare and isolate purine 3′-ketonucleosides have often resulted in elimination of the nucleoside base due to the high acidity of the H2′ protons.[214]
It was decided to follow a procedure described by Eisenhuth *et al.*, published in 2008, in which 3′-amino-2′-deoxyadenosine was synthesised via a 3′-ketonucleoside intermediate. However, in this case, the base had been protected with a benzoyl group.[215] The research carried out by this group found that low temperature

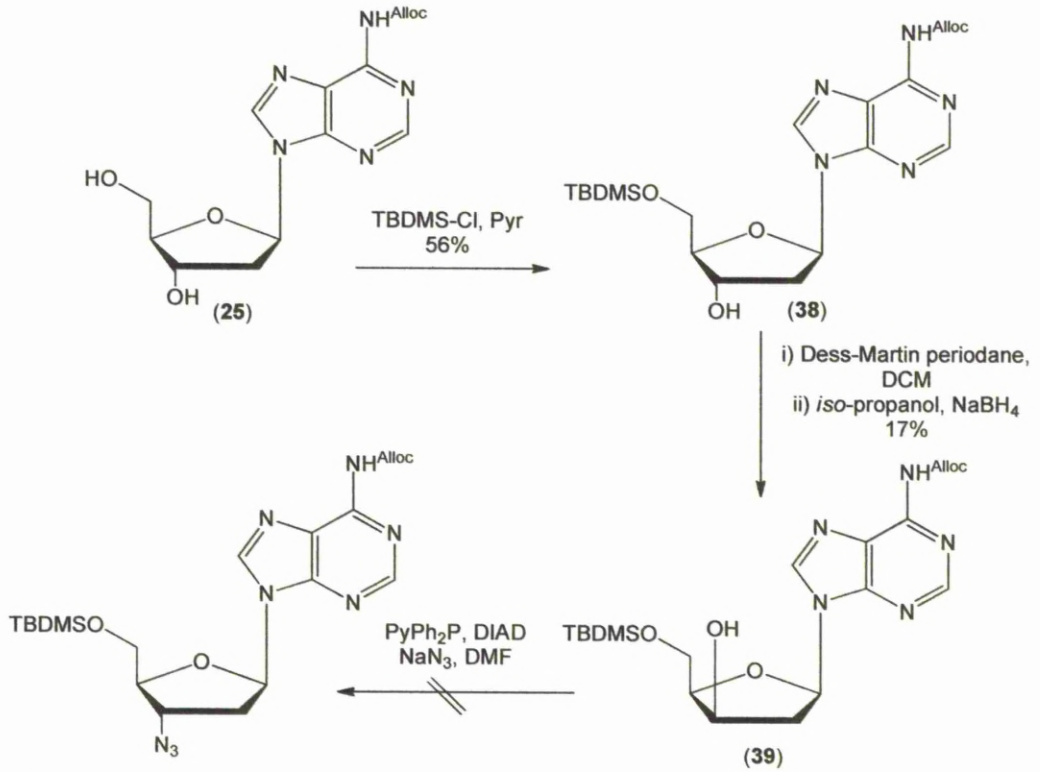
oxidation, followed by immediate reduction at -60°C gave a highly selective, high yielding route to a 3′-*xylo*-nucleoside with minimal depurination observed. Application of this route to the Alloc-protected 2′-deoxy-adenosine is shown in Scheme 2.14.

The first step in the synthesis was to introduce silyl protection at the 5′-OH group in order to avoid oxidation at this position. The group chosen was the bulky *t*-butyldimethylsilyl group (TBDMS). It was hoped that the size of this group would also assist with selectivity of the reduction later on in the synthesis, by partially blocking access to the top face of the molecule. Protection proceeded smoothly in 56% yield; some 3′-protection and 3′, 5′-diprotection was also observed which accounted for the major loss of yield in this reaction. Successful protection to give 5′-protected nucleoside (**38**) was confirmed by the appearance of characteristic singlets, in the ^{1}H NMR spectrum, at 0.01 ppm and 0.80 ppm. These correspond to the 6 protons of the methyl and 9 protons of the *t*-butyl group, respectively.

Oxidation and subsequent reduction of the 3′-position were carried out in succession in a 'one-pot' synthesis. Thus, the 5′-protected nucleoside (**38**) was treated with Dess-Martin periodane at 0°C and the progress of the reaction was easily be monitored by tlc. After 30 mins, all the starting material appeared to have been consumed. The excess oxidant was quenched by the addition of *iso*-propanol, which also served as the solvent for the reduction of the 3′-ketonucleoside. Fresh sodium borohydride was added to the reaction mixture at low temperature. However as the reaction proceeded, formation of a large amount of unidentified by-product was observed by tlc, as well as formation of the desired compound (**39**). After work-up and isolation by flash chromatography, it was found that this by-product was in fact the nucleoside base that had been eliminated from the molecule. The identity of this by-product was confirmed by NMR as only peaks corresponding to the Alloc protons and H2 and H8 of the adenine base were observed.

It was thought, that in this case, the hydride anion acted as a base extracting one of the acidic 2′-protons and inducing base elimination. Although multiple attempts at this reaction were carried out, the literature yield of 77% reported for the analogous *N*-benzoyl compound was never achieved. The maximum yield of (**39**) obtained in our hands was 17%. Although the starting material (**38**) and the 3′-*xylo*-nucleoside (**39**) were practically indistinguishable by tlc, a small upfield shift in the peak

assigned to H3', from 4.62 ppm to 4.48 ppm confirmed that the stereochemical inversion had indeed been successful.



Scheme 2.14: Inversion of stereochemistry at C3' using Dess-Martin periodane (TBDMS = *t*-butyldimethylsilyl)

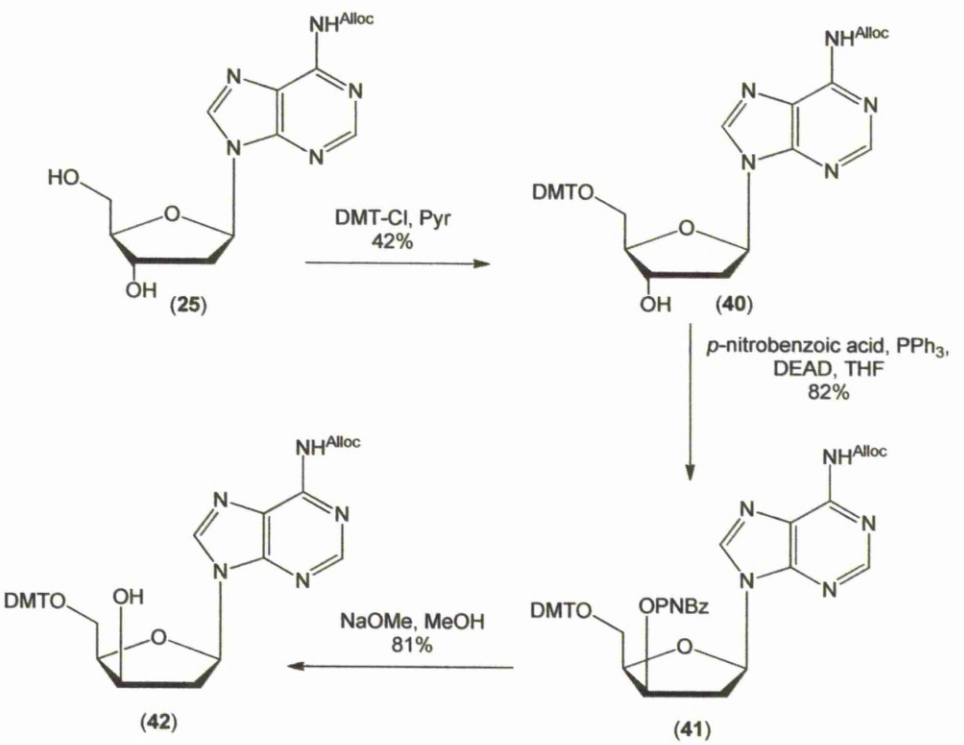
Although a poor yield had been obtained from the inversion step, it was decided to continue with the synthesis according to the paper and introduce the 3'-azide group *via* a Mitsunobu reaction which employed NaN_3 as the azide source. Although an attempt was made utilising the exact same conditions as reported by Eisenhuth *et al.*,[215] no reaction was observed here. This observation combined with the disappointing yield from the inversion step, brought us to the conclusion that no further time should be dedicated to this route.

2.4.4 Synthesis of adenosine-derived β-amino acid: Route C

Although the first two methods for inverting the C3' stereochemistry of 2'-deoxyadenosine proved to give unexpectedly poor results, a number of other methods still remained that we considered worth investigating including: hydride reduction of *lyxo*-epoxides,[216] regioselective 2'-*O*-tosylation of adenosine followed

by a 1,2-hydride shift [217] and inversion by Mitsunobu chemistry.[218] The latter of these methods appeared to offer the simplest route, without the need to obtain expensive reagents or complex starting materials. It was therefore decided to attempt this method first.

A publication by Lavandera *et al.* in which azide functionality had been introduced at both the 3′- and 5′-position was used as inspiration for the devised route[218] As with the previous two syntheses, the initial step in the reaction sequence was protection of the 5′-hydroxyl group. Initially, it was decided to follow the route as described by Lavandera *et al.* and use the DMT protecting group (Scheme 2.15). Protection of (**25**) proceeded smoothly, although in an unusually low yield of 42% after purification. However, this resulted in more than enough of the tritylated nucleoside (**40**) to carry through the synthetic pathway. Successful protection was confirmed by ^{1}H NMR spectroscopy through the appearance of characteristic aromatic peaks and a singlet at 3.77 ppm, corresponding to the 6 protons of the methoxy groups.

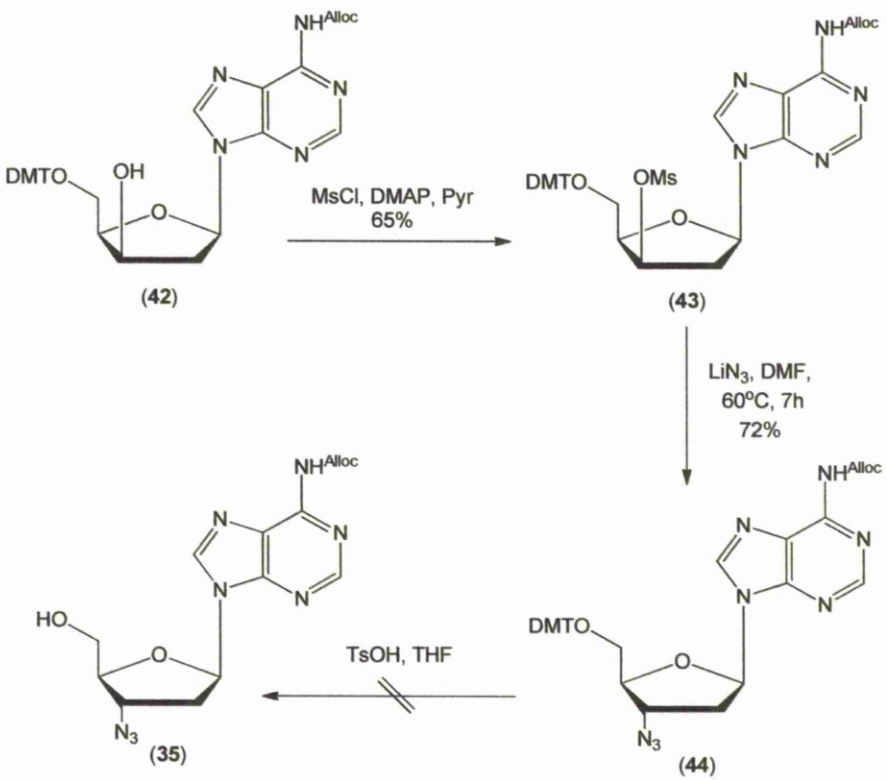


Scheme 2.15: Inversion of stereochemistry at C3′ using Mitsunobu conditions (PNBz = *p*-nitrobenzoyl).

Inversion of the 3′-hydroxyl group was carried out in two steps, the first of which was an S_N2 reaction using *p*-nitrobenzoic acid. Formation of the 3′-*p*-nitrobenzoate ester

(**41**) was carried out using standard Mitsunobu chemistry with the formation of triphenylphosphine oxide side product being the driving force for the reaction. The desired compound (**41**) was afforded in 82% yield, after purification by flash chromatography. The identity of (**41**) was confirmed by a downfield shift of the H3′ proton peak from 4.71 ppm to 5.89 ppm and the appearance of characteristic aromatic peaks corresponding to the *ortho* and *meta* protons of the *p*-nitrobenzene ring. A small loss of the DMT group was observed during this reaction which was not completely unexpected due to the presence of *p*-nitrobenzoic acid.

Cleavage of the ester moiety was carried out under basic conditions using sodium methoxide and afforded the 3′-*xylo*-2′-deoxyadenosine analogue (**42**) in 81% yield. Removal was confirmed by the loss of the aromatic peaks in the ^{1}H NMR spectrum corresponding to the *ortho* and *meta* protons of the *p*-nitrobenzene ring and a shift of the H3′ peak back upfield to 4.36 ppm.



Scheme 2.16: Attempted synthesis of 3′-azido-2′-deoxyadenosine analogue (**35**).

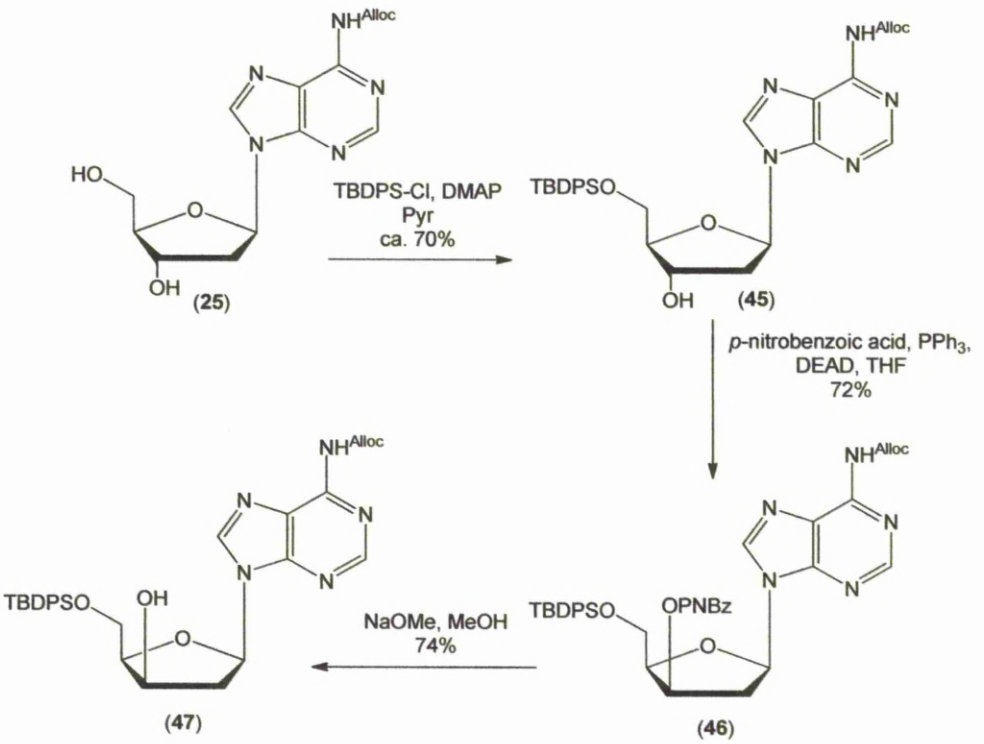
The newly inverted 3'-hydroxyl group was converted to a leaving group by treatment with methanesulfonyl chloride to allow for easy displacement by azide (Scheme 2.16). Mesylation proceeded in 65% yield and provided the sulfonate ester (**43**). Sucessful mesylation was confirmed by ^{1}H NMR with a downfield shift of the H3' peak from 4.36 ppm to 5.48 ppm and the appearance of a singlet at 3.07 ppm characteristic of the methyl protons of a SO_2CH_3 group. Displacement of the leaving group by azide was carried out in a simple S_N2 fashion by heating compound (**43**) in the presence of LiN_3. This afforded the desired 3'-azido-2'-deoxyadenosine derivative (**44**) in 72% yield. However, even after multiple washes and chromatographic purification, trace amounts of the reaction solvent (DMF) could still be seen in the ^{1}H NMR spectra. Formation of the azide compound was confirmed by a strong absorption in the infrared spectrum at 2102cm^{-1} characteristic of an azide.

Attempts to remove the DMT group from compound (**44**) under mildly acidic conditions resulted only in degradation of the nucleoside, presumably by cleavage of the glycosidic bond. A tlc of the reaction showed 3 distinct spots, one of which was the orange staining trityl group which runs with solvent front in the 5% methanol/DCM eluent system used. A second spot running slightly below the trityl group was found to be UV active but did not stain with anisaldehyde and a third lower running spot was not UV active but could be stained with anisaldehyde. It was thought that these were the base portion and the sugar portion of the nucleoside, respectively.

After such encouraging results with the Mitsunobu method of inversion, it was decided to re-think the 5'-hydroxyl protection and examine groups that did not require acidic conditions for their removal. After a review of protecting groups used in nucleoside chemistry, the decision was made to use a *t*-butyldiphenylsilyl (TBDPS) group for protection of the 5'-position. This group was chosen due to its high selectivity for primary hydroxyl groups over secondary and its ease of removal under standard conditions with TBAF. It was hoped that the combination of these factors would help to maximise the overall yield of the synthetic pathway.

Protection of the 5'-hydroxyl group of (**25**) with TBDPS chloride to give (**45**) proceeded without any discernable observation of 3'-protection or 3', 5'-diprotection. However, due to contamination of the starting material (**25**) with tetrabutylammonium salts, only an approximation of the yield (70%) could be established (Scheme 2.17). Silylation was confirmed by the appearance of a singlet

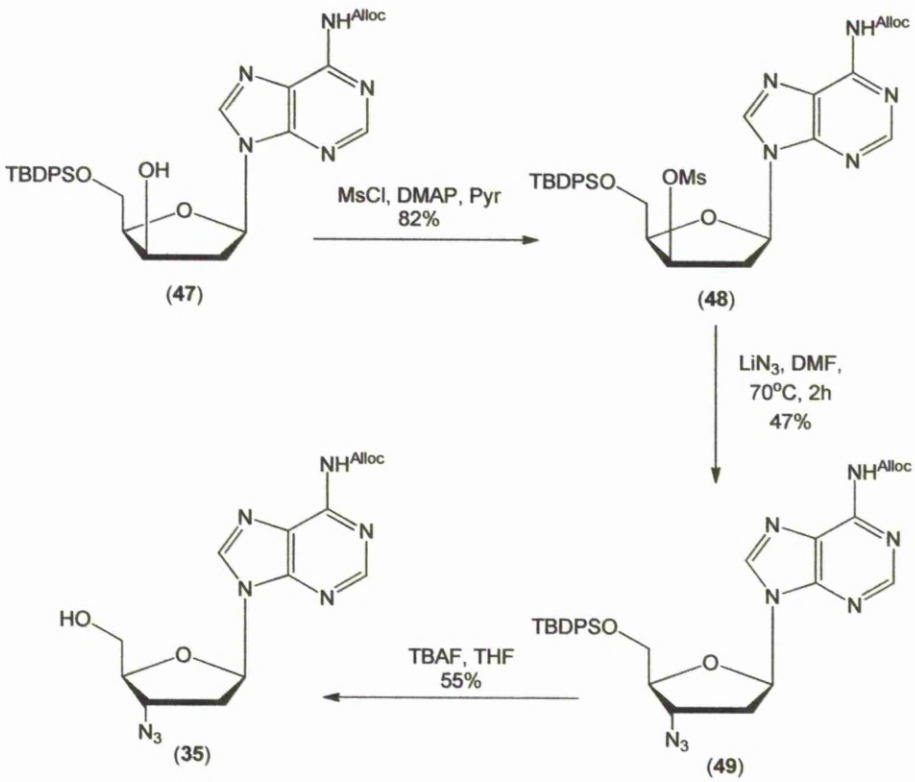
at the low chemical shift of 1.04 ppm in the ^{1}H NMR spectrum, corresponding to the *t*-butyl protons; a signal at this position is characteristic of alkyl groups attached directly to a silicon atom. Further to this, aromatic peaks were observed corresponding to 10 protons. Introduction of the *p*-nitrobenzoate ester and cleavage were carried out in the same manner as previously described and afforded the 5′-TBDPS-3′-*xylo* nucleoside (**47**) in 54% yield over 2 steps. Although the individual yields for these steps were slightly lower than the previous attempt with the analogous DMT protected derivatives, the yields from this method of inversion were still far superior to those that had been obtained from the Herdewijn and Eisenhuth routes.



Scheme 2.17: Inversion of stereochemistry at C3′ using Mitsunobu conditions (TBDPS= *t*-butyldiphenylsilyl).

Mesylation of the 3′-*xylo*-nucleoside (**47**) was carried out in the same manner as previously described and proceeded smoothly to give sulfonate ester (**48**) in an 82% yield (Scheme 2.18). This slight improvement in yield can possibly be attributed to the greater stability of the TBDPS group to the reaction conditions, when compared to the DMT group. Displacement of the mesylate group with azide, however, did not proceed as expected. Subjection of a solution of sulfonate ester (**48**) in DMF to LiN_3 at 60°C for 7 hours resulted in not only introduction of the azido group, but also

removal of the allyloxycarbonyl protection from the base. This problem was overcome by carrying out the reaction at a slightly elevated temperature (70°C rather than 60°C) over a shorter time period (2 hours rather than 7 hours). Isolation of the 3′-azido nucleoside (**49**) proved difficult due to the formation of emulsions during the extraction process and problems were encountered in isolating pure product from the reaction solvent DMF. Consequently, the yields from this step were greatly reduced. A 42% yield of the 3′-azido analogue (**49**) was isolated in high purity and successful reaction was confirmed by the presence of an azide stretch at 2102cm^{-1} in the infrared spectrum. In addition, an upfield shift of the H3′ peak from 5.45 ppm to 4.49 ppm in the ^{1}H NMR spectrum was observed. The presence of the Alloc group was also confirmed by ^{1}H NMR spectroscopy by a characteristic multiplet at 5.91 ppm corresponding to the alkenyl CH and two apparent double quartets at 5.21 ppm and 5.33 ppm corresponding to the terminal alkenyl protons.

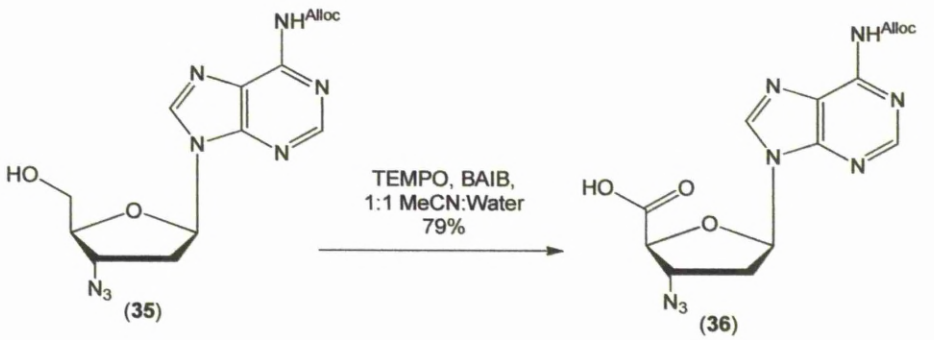


Scheme 2.18: Synthesis of 3′-azido-2′-deoxyadenosine analogue (**35**).

Removal of the silyl protection was carried out by treatment of (**49**) with TBAF to afford the 3′-azido-2′-deoxyadenosine derivative (**35**) in 55% yield. Deprotection was

confirmed by the loss of the peak at 0.98 ppm in the ^{1}H NMR spectrum of the starting material (**49**) which had been attributed to the 9 protons of the *t*-butyl group.

It was decided at this point, in the interest of time and due to the problems experienced previously with the Fmoc protection step, to use the azide itself as a masked amine for the synthesis of the oligomers. This not only shortened the synthesis but also reflects the fact that good yields can be obtained for the reduction of the azide, under a variety of conditions. Therefore the last remaining step in this synthesis was oxidation of the 5′-position (Scheme 2.19).



Scheme 2.19: Oxidation to produce final product (**36**).

Oxidation of the 5′-position to give the carboxylic acid using the standard TEMPO/BAIB conditions, previously described, gave the desired azido acid (**36**) in a pleasing, 79% yield. Isolation of pure compound was carried out by dissolution of the crude reaction residue in the minimum amount of DCM and precipitation of the desired acid (**36**) using diethyl ether. Successful oxidation was confirmed by the loss of the H5′ protons seen at 3.82 ppm and 4.03 ppm in the ^{1}H NMR spectrum of the starting material (**35**).

Although the protected amino acid derived from 2′-deoxyadenosine was not obtained as originally designed, the 3′-azido-2′-deoxyadenosine-5′-carboxylic acid compound (**36**) can still be utilised as a useful building block for the synthesis of β-peptides in solution. Coupling of this monomer will be discussed in Chapter 3.

2.5 Conclusions

Fmoc amino acids of thymidine and 5-methyl-2′-cytidine were both successfully synthesised from the parent nucleoside thymidine utilising the widely studied, pharmaceutically active compound AZT as a common intermediate. Although there still remains room for improvement in the synthesis of the 5-methyl-2′-cytidine amino acid, in terms of introduction of a 3′-*N*-protecting group, the route devised provided the compound in sufficient quantities and purity for coupling studies to be carried out. Further to this, a second thymidine derived monomer with benzhydryl protection of the 5′-carboxylic acid group has been successfully synthesised and this monomer was expected to be better suited to synthesis of β-peptides in solution.

The synthesis of an Fmoc amino acid derived from 2′-deoxyadenosine proved problematic with difficulties encountered in both the inversion of stereochemistry at C3′ and introduction of the azide moiety. However, these issues were overcome after investigating a variety of synthetic routes and protecting groups. Although the Fmoc-protected amino acid was not prepared, an alternative 3′-azido-5′-carboxylic acid monomer was synthesised successfully which could be used instead in the solution synthesis of the β-peptides.

It should be noted that a significant potential problem with these β-amino acids, particularly those derived from 2′-deoxyadenosine, is the lability of the glycosidic bond. For this reason, it is important to avoid conditions that are even mildly acidic, and this should be taken into consideration in any subsequent work performed on 2′-deoxyguanosine.

Chapter 3: Results and Discussion 2

Synthesis of β-peptides

3.1 Peptide synthesis: Project aims

The next goal in this project was to couple the nucleoside derived β-amino acid monomers discussed in chapter 2, utilising amide bond formation, to synthesise short β-peptides. Previous research within the group has shown that secondary structures can be observed in peptides as short as four residues in length and that this can be studied in detail by NMR.[153] However, the oligomers synthesised previously from the thymidine derived β-amino acid possessed an amide functionality at what would be described as the '*C*-terminus' of the peptide (blue, Figure 3.1). We wanted to extend this study and examine the feasibility of obtaining oligomers that would have a carboxylic acid *C*-terminus (pink, Figure 3.1) so that both secondary structure and cyclisation studies could be carried out.

Assembly of peptides can be carried out by two different methods: solid-phase peptide synthesis (SPPS) and solution peptide synthesis. Both methods were applied to the synthesis of these β-peptides and each will be discussed in turn within this chapter.

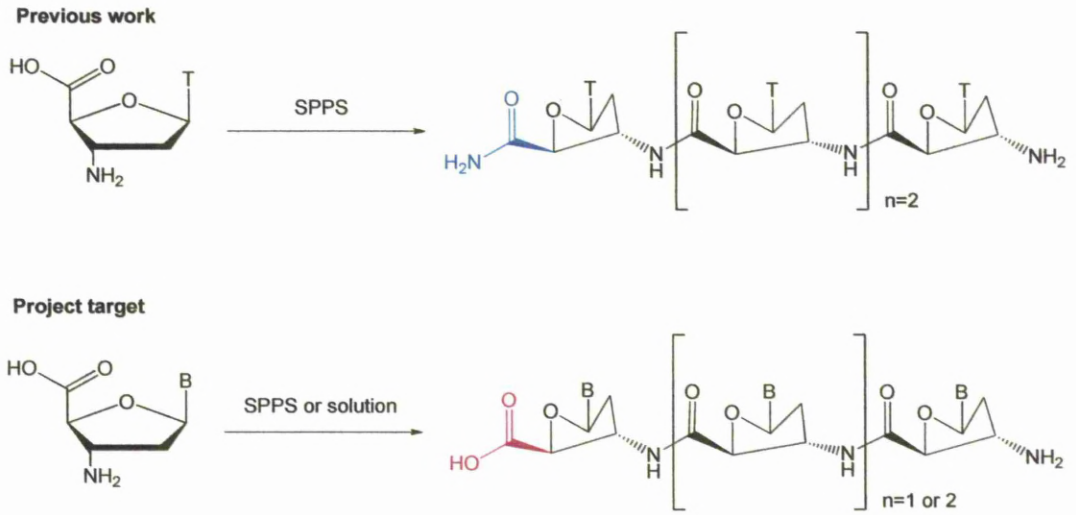


Figure 3.1: Nucleoside β-peptides: Previous work (top) and project target (bottom) (T= thymine and B=nucleobase).

3.2 Solid-phase peptide synthesis (SPPS)

3.2.1 A brief overview of SPPS

The original method of solid-phase peptide synthesis was developed and reported by Merrifield in 1963.[219] The work exploited the use of an insoluble, yet porous, solid support containing a chlorobenzyl linker to which a Boc-protected amino acid could be covalently attached. The Boc group could then be removed from the first amino acid residue under acidic conditions and subsequently coupled to a second residue by amide bond formation. This simple, yet effective, procedure revolutionised peptide synthesis and began the development of the methods we use today.

An improved method developed by Merrifield utilising Boc protection of the α-amino function and side chain benzoyl protection of all other nucleophilic centres, had a number of disadvantages associated with the use of strongly acidic conditions required to deprotect the amino acids and remove the peptide from the resin.[220] The use of protecting groups which can all be cleaved by similar conditions lead to concomitant loss of a small proportion of side chain protecting groups and degradation of a small amount peptide-resin linkage at each stage in the cycle. Loss of side-chain protection can result in the liberation of reactive groups which can participate in undesired side reactions, this results in failure sequences and consequently a reduction in final product purity. Cleavage of a small amount of peptide at each stage also results in a further cumulative loss of yield.

With these problems becoming apparent, a lot of the work into improved methods of solid-phase peptide synthesis has focussed on the development of mild conditions for the removal of the peptide from the resin and the development of orthogonal protecting group strategies to minimise peptide and side-chain protecting group cleavage.

The publication of the Fmoc group as a base-labile *N*-terminal protecting group for amino acids, in 1972,[221] promoted a huge leap in the development of mild acid-free deprotection conditions for solid-phase peptide synthesis. This group can be removed rapidly by treatment with an organic secondary base without any degradation of peptide bonds. For this reason it became, and still remains, an extremely popular terminal protecting group in peptide chemistry.

As a consequence of the discovery of base-labile protecting groups, such as the Fmoc group, that can be used in solid phase peptide synthesis a number of resin

linkers have been developed that require only extremely mild acidic conditions to liberate the peptide. Many of these are based upon Merrifield's original design and release the peptide with either a *C*-terminal carboxylic acid or amide.

Commonly used resins for the synthesis of peptides with a carboxylic acid terminus include the Wang resin,[222] Rink acid resin[223] and HMPB-BHA derived resins.[224] The most commonly used resin for the generation of peptides with an amide terminus is the Sieber amide resin[225]. Removal of peptides from all of the resins discussed here requires only a 1-5% solution of TFA, which, compared to the conditions required to cleave peptides from the Merrifield resin (HF), offers less cause for concern over undesirable removal of side chain protection or peptide degradation.

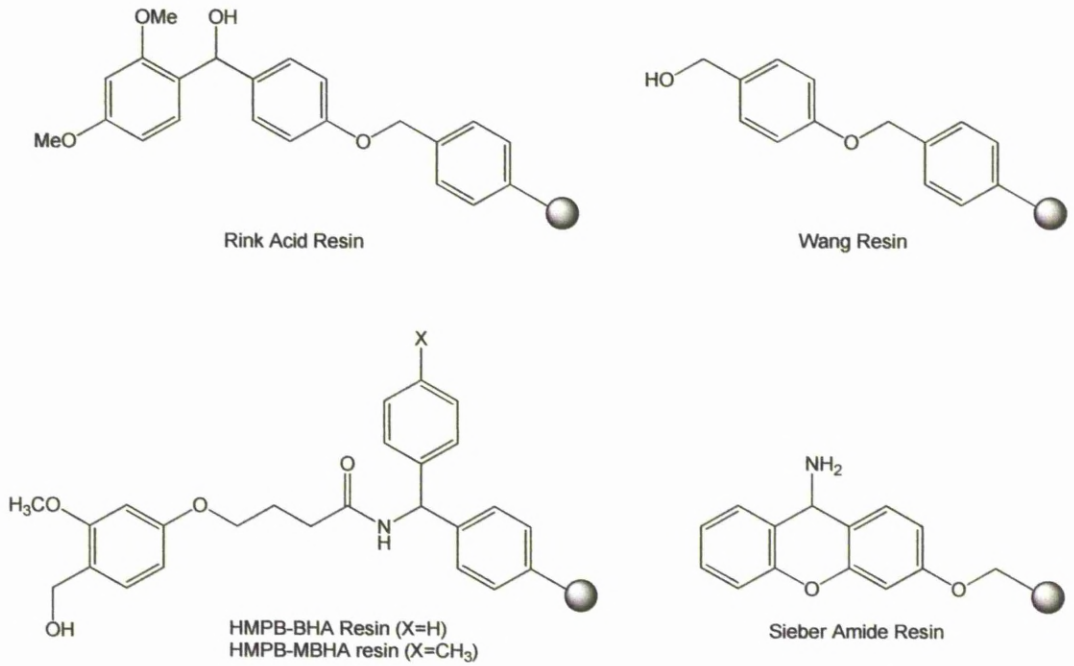


Figure 3.2: Resins used in SPPS.

3.2.2 Peptide Synthesis by SPPS

Previous work within our group had focussed on the synthesis of short β-peptides of the thymidine derived β-amino acid synthesised on the Sieber amide resin. This resin was chosen due to the high acid-lability of the xanthen-9-amine linker which can be cleaved with exposure to a 1% TFA in DCM solution. As alluded to in Chapter 2, the glycosidic bond in nucleosides, particularly the purines, is extremely acid labile so exposure to acidic conditions must be kept to a minimum in order to

maximise yield and purity of the peptides. For initial studies of secondary structure, the fact that this resin would result in an amide terminus was of no great concern. However, we wished to now carry out further studies of these peptides and assess their potential to form macrocycles. This required the synthesis of the same short β-peptide structures but with a carboxylic acid terminus rather than an amide.

Previously within our group, a study had been carried out on three resins that could potentially provide peptides with a carboxylic acid terminus. These were the Wang resin, Rink acid resin and HMPB-BHA resin (Figure 3.2). These resins were of particular interest as the linkers require the mildest acidic conditions for removal of the peptide. It was found that the best results (60% coupling efficiency) were provided by the HMPB-BHA resin (which has a dialkyloxybenzyl linker) with a coupling reagent combination of PyBOP/DIPEA/DMAc. However, immediately after this study had been carried out, the HMPB-BHA resin was discontinued and replaced by a similar HMPB-MBHA resin (Figure 3.2). Even though the change to the linker was relatively small, the coupling conditions required reassessment in order to provide the best initial coupling yield possible.

There are numerous permutations of coupling reagents that can be used in SPPS. However, we chose to test the new resin with the 3 combinations of reagents that had given us the best coupling yields in the past. The results of this study are outlined in Table 3.1.

Activator/solvent/base	**Coupling efficiency***
PyBOP/DMAc/DIPEA	No coupling
HBTU/DMAc/DIPEA	62%
HBTU/DMAc/DIPA	48%

Table 3.1: Coupling efficiency of the thymidine derived β-amino acid to HMPB-MBHA resin. *Coupling efficiency was determined by spectrophotometric measurement of the dibenzofulvene released upon Fmoc deprotection (see experimental section for details).

The results of this study indicated that the most suitable coupling combination for attachment of the first nucleoside residue to the resin had changed from a system utilising PyBOP/DMAc to a system requiring HBTU/DMAc. It is not apparent what effect is responsible for change in coupling efficiency with different reagents. A

subsequent second coupling carried out under the same reaction conditions was found to be quantitative.

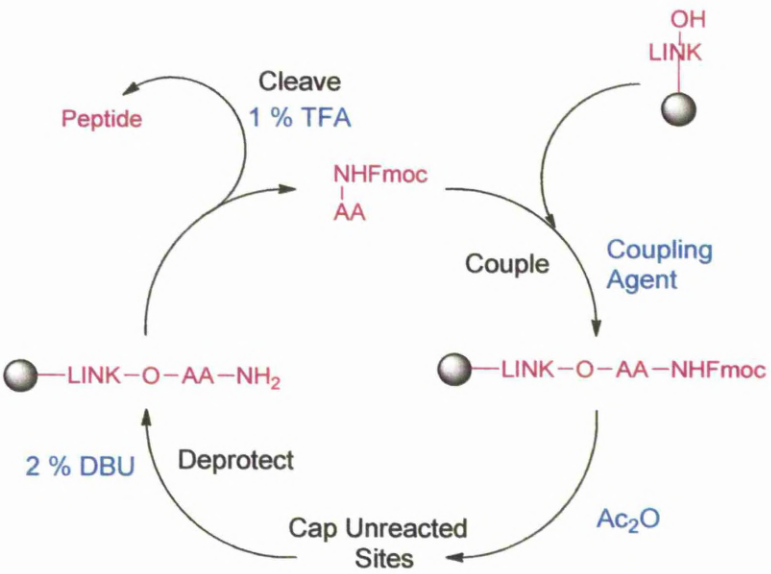
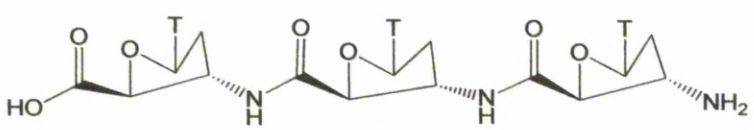


Figure 3.3: SPPS cycle

With the identification of suitable coupling conditions for the peptide synthesis complete, assembling of a peptide three residues in length could now begin (**50**, Figure 3.4). The resin was swollen overnight in DCM and then drained and washed with DMAc. The thymidine β-amino acid was pre-activated for 60 seconds with the coupling reagent system and subsequently added to the resin bed which was then agitated for five hours. The excess reagents were then rinsed from the resin and the process of capping un-reacted sites, followed by removal of the Fmoc group, were performed to enable the second coupling to take place. Once the peptide of desired length had been assembled the resin bed was rinsed with DMAc, DCM and methanol and dried under vacuum overnight. Re-swelling of the resin allowed for acidic cleavage of the peptide from the linker with a 1% TFA solution in DCM (this process is outlined in Figure 3.3). The liquors were then removed from the peptide and analysis of the crude mixture carried out by mass spectroscopy.



(50)

Figure 3.4: peptide assembled by SPPS.

It was found that the peptide had indeed been formed. However, the presence of a lot of other fragment peaks prompted us to investigate the purity of the oligomer by RP-HPLC (Figure 3.5). The HPLC trace of the peptide showed a pattern of 5 major peaks, which were isolated individually and analysed by mass spectroscopy.

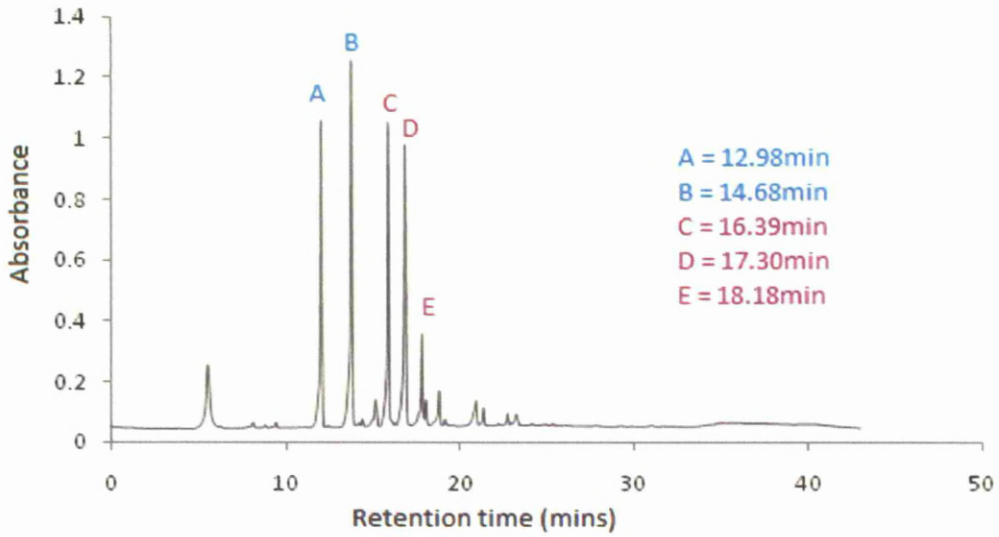


Figure 3.5: RP-HPLC trace of the peptide (**50,** for conditions refer to experimental)

The results from electrospray mass spectrometry confirmed that the first two major peaks (A +B) were failure sequences of the peptide. However, the final 3 major peaks (C, D +E) all gave the same molecular ion which was consistent with the full length trimer (mass 728). The fact that the trimer appeared as multiple peaks in the HPLC indicated that perhaps the peptide may have been forming distinct secondary structures in solution. Attempts to break down the suspected secondary structures and coalesce the 3 peaks by using a HPLC column heater resulted in partial success with the formation of two broadened peaks. Unfortunately, though, a single peak which would have aided purification and analysis was not achieved.

Due to the difficulties encountered with the purification of this short, fully deprotected peptide, this method of peptide synthesis was abandoned. This decision was also further justified when the synthesis of the 2′-deoxyadenosine β-amino acid monomer, which was carried out in parallel with this work, revealed the extremely sensitive nature of the glycosidic bond of these modified nucleosides, such that they would be incompatable with even very acid labile linkers. The combination of these factors and the fact that we required only short oligomers, led us to the decision to explore peptide synthesis in solution as an alternative.

3.3 Peptide synthesis in solution

3.3.1 Amide bond formation in solution

The synthesis of peptides in solution can be carried out utilising simple amide bond forming reactions. This method does have some advantages over SPPS; it can be carried out on larger scales and the intermediates can be isolated and analysed at each step. However, it also has some distinct disadvantages which include a longer and more cumbersome synthesis and problems associated with solubility of larger peptides due to the neutral nature of the backbone. As we required short peptides for this study, we felt that, in this case, the advantages of this method outweighed the disadvantages.

The formation of an amide bond is one of the most important reactions in organic chemistry and many different strategies have been devised to form these bonds between carboxylic acids and amines. A common problem of amide bond forming reactions in peptide chemistry is the possibility of amino acid racemisation once the acid moiety has been activated. This racemisation is induced by intramolecular oxazolone formation (Figure 3.6) and can cause severe problems when peptides are synthesised for biological purposes. However, as our nucleoside derived β-amino acids contain a conformationally constrained furan ring system, we reasoned that this potential side reaction was unlikely to occur.

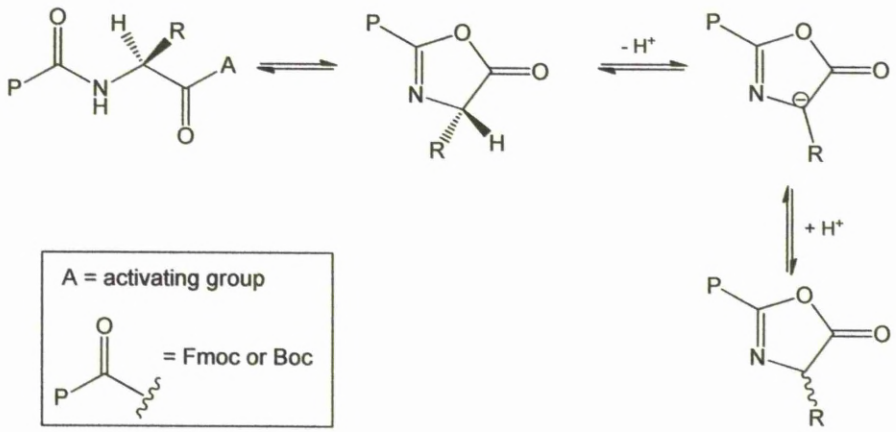
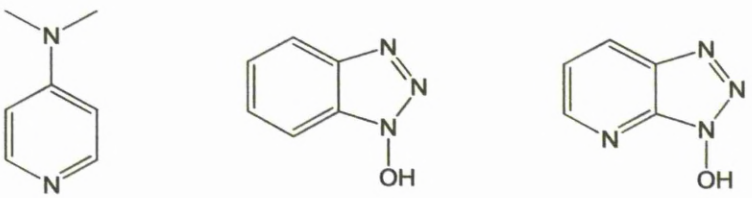


Figure 3.6: Intramolecular oxazolone formation in activated amino acids.

3.3.2 Coupling reagents

Some of the most commonly used reagents for amide bond formation in organic synthesis include: carbodiimide-type, phosphonium-type and uronium-type reagents.[226]

Within the area of peptide and oligonucleotide chemistry the most common carbodiimide-type coupling reagents employed are dicyclohexylcarbodiimide (DCC),[227] di*iso*propylcarbodiimide (DIC)[228] and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).[229] These reagents are usually used in combination with a nucleophilic catalyst such as *N,N*-dimethylaminopyridine (DMAP), 1-*N*-hydroxybenzotriazole (HOBt) or 7-aza-1-*N*-hydroxybenzotriazole (HOAt) (Figure 3.7).



DMAP HOBt HOAt

Figure 3.7: Nucleophilic catalysts used in carbodiimide coupling reaction.

The carbodiimide reagent reacts with the acid moiety and converts this to an acyl *iso*urea species which can then undergo nucleophilic attack by the catalyst to form either an activated ester, in the case of HOBt and HOAt, or a reactive amide in the case of DMAP. In the absence of a nucleophilic catalyst, the acyl *iso*urea species can undergo an undesired rearrangement to form the *N*-acyl urea species which renders the acid unable to couple with its partner amine (Figure 3.8). The addition of the nucleophilic catalyst not only prevents this rearrangement from occurring but it also creates an extremely reactive ester which, when treated immediately with the partner amine, reacts at such a rate that almost no racemisation is usually observed.

A common problem with the use of carbodiimide coupling reagents is the removal of the urea by-products. However, recent years have seen the development of newer coupling agents, such as DIC which have enhanced solubility in organic solvents. These newer coupling agents are more suited to solid-phase synthesis as they can be easily removed through a series of washes.[230]

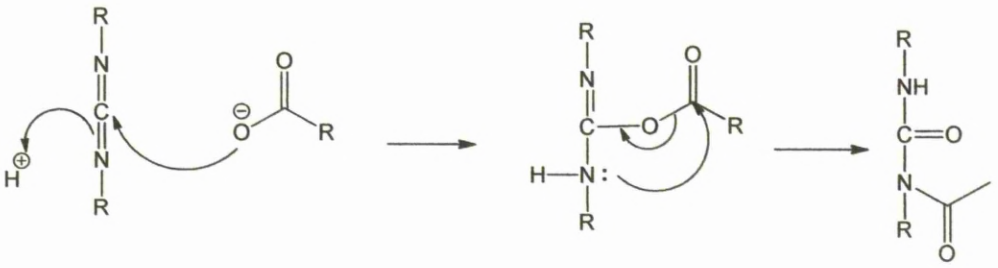


Figure 3.8: Intramolecular rearrangement by carbodiimide activated acids.

Phosphonium-type peptide coupling reagents were first developed in the 1970's by Castro *et al.*.[231,232] These early coupling agents produced a large amount of racemised material. However, the discovery of HOBt as a racemisation suppressor[227] led to the development of a new generation of coupling reagents which remain in frequent use today. The first of these reagents was benzotriazol-1-yloxy*tris*(dimethyl-amino)-phosphonium hexafluorophosphate (BOP)[233] (Figure 3.9). At the time of its development, BOP was a popular coupling agent used throughout peptide and organic chemistry.[234] However, the realisation that the by-product from this reagent (hexamethylphosphoramide, HMPA) was a potent human carcinogen lead to further modifications to produce new compounds that did not generate toxic by-products. The modifications to the BOP reagent resulted in the reagents that are known today as benzotriazol-1-yloxytri(pyrrolidino)-phosphonium hexafluorophosphate (PyBOP)[235,236] and the HOAt derivative [(7-azabenzotriazol-1-yl)oxy]*tris*-(pyrrolidino)-phosphonium hexafluorophosphate (PyAOP).[237,238]

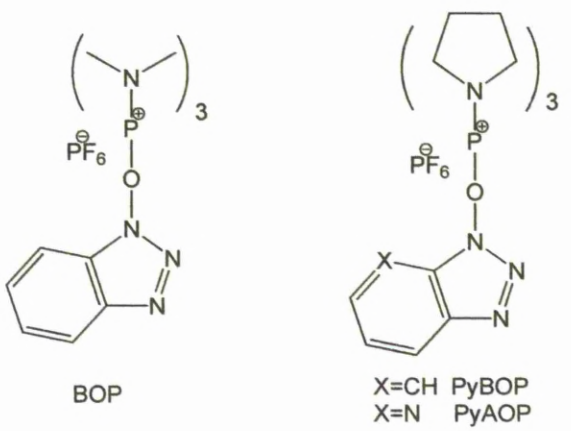
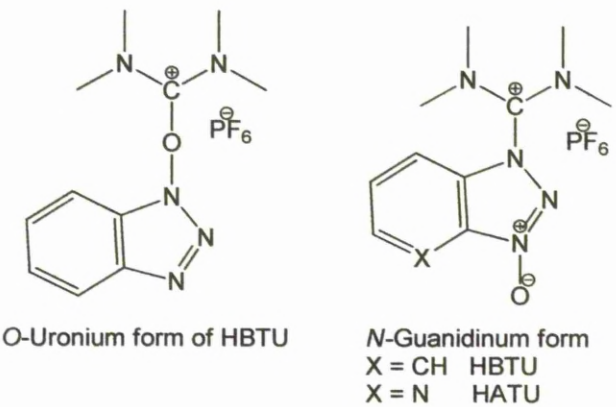


Figure 3.9: Phosphonium type coupling agents.

The proposed mechanism of these phosphonium-type reagents proceeds initially with deprotonation of the acid which then attacks the positively-charged phosphorus

causing elimination of the OBt (or OAt) portion of the molecule. This OBt anion then acts as a nucleophile attacking the carbonyl group of the newly formed phosphorus intermediate causing elimination of a phosphorus oxide by-product. The activated ester can then react with an amine to form an amide bond.

Uronium-type coupling reagents were first introduced in 1978 by Gross *et al.*[239] This group developed the reagent *O*-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) which remains in wide use today.[239,240] It was initially thought that this reagent existed as the *O*-uronium salt. However, recent studies have shown that the active form of the compound is actually the *N*-guanidinium salt[241] (Figure 3.10). Along with HBTU, the HOAt derived compound *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) is also frequently used.[237,242] Although many uronium-derived coupling agents exist, HBTU and HATU are two of the most popular due to their commercial availabilities.



O-Uronium form of HBTU

N-Guanidinum form
X = CH HBTU
X = N HATU

Figure 3.10: Uronium-type coupling reagents.

It is thought the mechanism of this class of reagents proceeds in a similar manner to that of the phosphonium-type reagents. Formation of a guanidinium ester of the carboxylic acid is subsequently followed by generation of an active OBt (or OAt) ester. This ester can then undergo nucleophilic attack by the partner amine generating the amide bond.

In general, the uronium based reagents tend to be more stable than the phosphonium based reagents and, for both classes of compound, the HOAt derivatives are found to be more reactive than their HOBt derivatives.[243]

3.3.3 Synthesis of peptides in solution

After reviewing literature in the area, it was decided to follow a procedure devised by Chandrasekhar *et al.* in which thymidine derived β-amino acid monomers, very similar to our own, were coupled together to make a very short peptide of four units in length.[157] The procedure utilised a HBTU/HOBt combination of coupling reagents to connect two thymidine-derived β-amino acid units through an amide bond. This nucleoside dimer was then used to synthesise a trimer and tetramer.

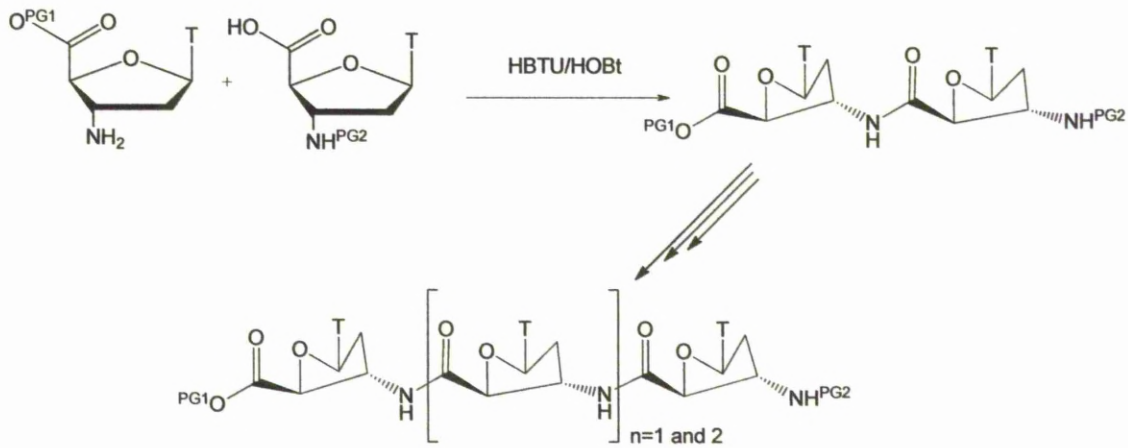
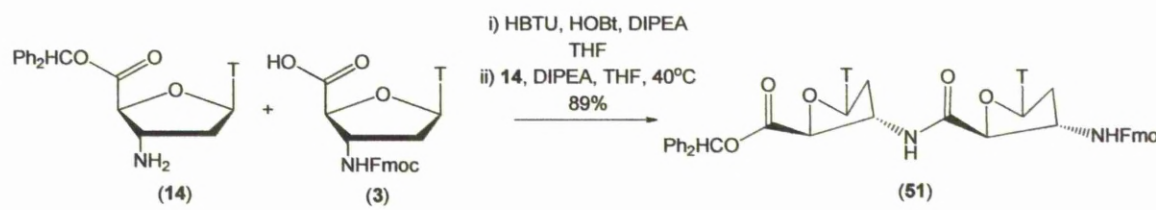


Figure 3.11: A general synthesis of short peptides in solution using HBTU/HOBt (PG1 and PG2 = protecting groups and T= thymine)

We wished to adopt the same strategy (see Figure 3.11) for the synthesis of peptides in solution built up of the nucleoside derived β-amino acid monomers discussed in chapter 2. Initially, it was decided to explore the preparation of homo-oligomers of the thymidine derived β-amino acids to assess the validity of this coupling method. Thus, the 5′-benzhydryl protected thymidine monomer (**14**) and 3′-Fmoc protected thymidine monomer (**3**) were coupled in solution to give fully protected 3′→5′ amide linked dimer (**51**).

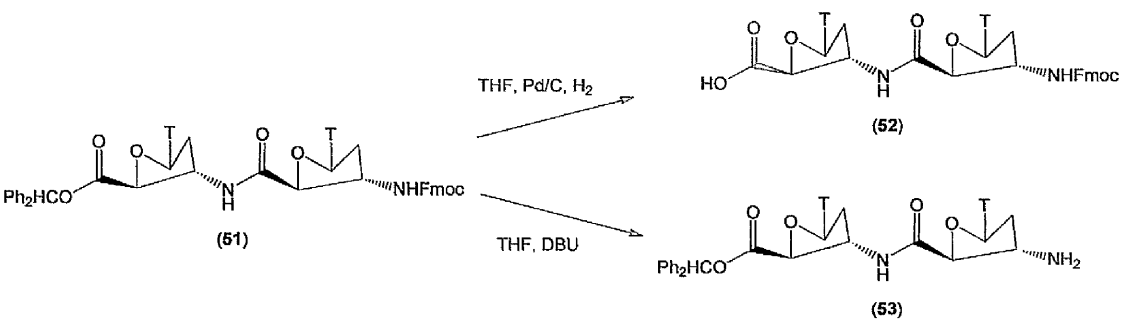


Scheme 3.1: Coupling of thymidine derived β-amino acid monomers to make a fully protected dimer.

This reaction was carried out by first activating the carboxylic acid moiety of amino acid (**3**) by treatment with a combination of HBTU/HOBt/DIPEA over a period of 2 hours. Addition of the second amino acid (**14**) in the presence of base enabled nucleophilic attack of the free 3'-amine on the newly activated acid and formation of the amide bond. In the publication by Chandrasekhar, this step reportedly took 3 days at ambient conditions and produced a 60% yield. However, we found that gentle heating at 40°C allowed our reaction to go to near completion in high yield (89% after purification by flash chromatography) within 1 hour. It is thought that this increase in yield and reduction in reaction time might be attributed to the enhanced solubility of our nucleoside derived β-amino acids due to a different protecting group strategy.

Formation of the fully protected dimer (**51**) was confirmed by ^{1}H NMR spectroscopy through the presence of two singlet peaks at 1.75 ppm and 1.91 ppm integrating to 3 protons each, corresponding to the two thymine methyl groups. The observation of four NH protons, two at 9.22 ppm and 10.03 ppm corresponding to the two amide NH groups and the remaining two at 13.09 ppm and 13.12 ppm corresponding to the two imine protons of the thymine base, provided additional support that the dimer (**51**) had indeed successfully formed.

The next stage in the synthesis was to generate two different dimers; one with a free acid moiety at the 5'-end (i.e **52**) and one with a free amine at the 3'-end (i.e. **53**, see Scheme 3.2). It was hoped that these two dimers could then be coupled together under the same conditions discussed previously to generate a fully protected peptide four units in length. Further to this, it was thought a dimer with a free 3'-amine moiety could be coupled to an additional single thymidine β-amino acid (**3**, Scheme 3.1) to generate a fully protected peptide three units in length.



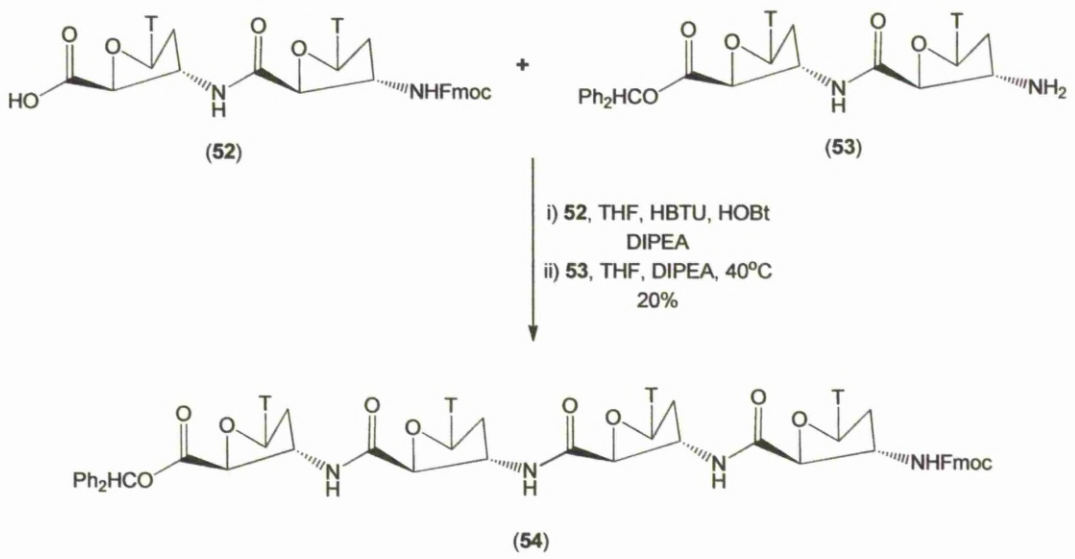
Scheme 3.2: Removal of protecting groups from the 3' and 5' end of nucleoside dimer (**51**).

Removal of the benzhydryl group from the carboxylic acid moiety was carried out under standard hydrogenation conditions. After 72 hours, the 5'-carboxylic acid dimer (**52**) was obtained. Purification by standard flash chromatography could not be carried out on this compound due to its high polarity. Instead a partial purification of the crude material was carried out by partitioning between methanol and hexane. The polar peptide was retained in the methanol and the less polar impurities were transferred into the hexane. As a consequence of the partially crude nature of the material, analysis by NMR was found to be extremely difficult. However, a good high resolution mass spectrum was obtained finding as mass of 737.2171 compared to the calculated mass of 737.2183.

Removal of the Fmoc group from the 3'-end of the fully protected dimer (**51**), to generate the dipeptide (**53**) with a free amine moiety, was carried out in THF using 1, 8-diazabicyclo-[5.4.0] undec-7-ene (DBU). Again the compound proved to be extremely polar so a partial purification was carried out as discussed previously by utilising a methanol/hexane extraction. A good high resolution mass spectrum of this compound was also obtained with a mass of 659.2490 being observed compared to the calculated mass of 659.2466.

Coupling of 5'-deprotected dimer (**52**) and 3'-deprotected dimer (**53**) was carried out under the same conditions as discussed previously in this chapter (Scheme 3.3). The carboxylic acid of dimer (**52**) was first activated using HBTU/HOBt/DIPEA followed by subsequent addition of 3'-amino dimer (**53**) in a mixture of THF and di*iso*propylamine. Once the reaction was judged to be complete, as determined by tlc, the reaction solvents were removed and the tetramer (**54**) was precipitated from methanol as a pale brown amorphous solid in 20% yield. The reason for the poor yield, in this case, is unknown as analysis of the reaction liquors determined that no residual starting materials were present.

The complexity of this molecule meant that analysis by NMR methods would prove to be difficult. An effort was made to obtain a ^{1}H NMR spectrum of this compound. However, it was found that, due to overlapping peaks, the structure could not be analysed with complete certainty and definitive peak assignments could not be made. Analysis of this compound by mass spectrometry also proved difficult due the reluctance of the molecule to ionise. However a low resolution mass spectrum could be obtained which gave the expected mass of the molecule (1377.2).

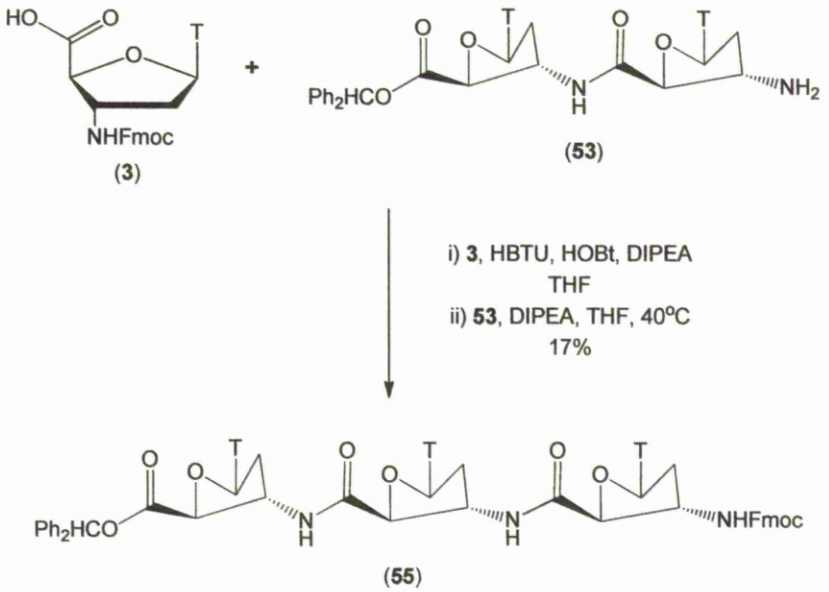


Scheme 3.3: coupling of dimers (**52**) and (**53**) to generate tetramer peptide (**54**).

In addition to the thymidine derived tetramer (**54**), a shorter fully protected peptide three residues in length (**55**) was also synthesised. Coupling of the 3′-amino dimer (**53**) to 5′-carboxylic acid monomer (**3**) using the HBTU/HOBt/DIPEA conditions described previously for the synthesis of the tetramer, afforded the trimer in a 17% yield (Scheme 3.4). Again the peptide was precipitated from methanol as a pale brown amorphous solid and the fate of the rest of the material could not be accounted for.

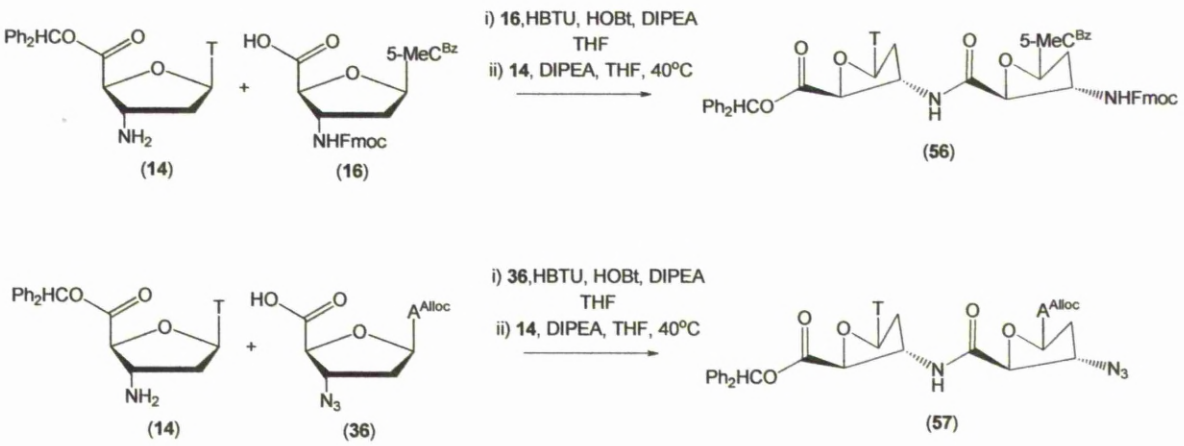
Analysis of the fully protected tripeptide (**55**) by electrospray mass spectrometry proved to be problematic in this case. The trimer proved to be difficult to ionise and the high cone voltage required to induce ionisation also resulted in the loss of the benzhydryl group from the 5′-end of the molecule. The only observed mass ion peak in this case corresponded to the mass of the peptide less a $C_{13}H_{10}$ fragment, giving 974.1. It is highly probable that the uncharged and fully protected trimer would be better suited to anaylsis by the MALDI-TOF technique rather than electrospray mass spectrometry.

Analysis of the tripeptide (**55**) by ^{1}H NMR again proved difficult due to overlapping peaks, but the spectrum was consistent with that expected for the product trimer. In particular, the aromatic protons of the benzhydryl and Fmoc groups could be seen over a range from 7.20 to7.70ppm and three singlet peaks corresponding to the thymine methyl groups were observed at 1.07 ppm, 1.11 ppm and 1.19 ppm.



Scheme 3.4: Coupling of 5′-carboxylic acid monomer (**3**) to 3′-amino dimer (**53**) to generate peptide trimer (**55**).

Due to the successful coupling of the thymidine derived β-amino acid monomers it was decided to investigate the coupling of the 5-methyl-2′-deoxycytidine derived monomer (**16**) and the 2′-deoxyadenosine derived monomer (**36**) to the 5′-protected thymidine derived amino acid (**14**). These reactions were carried out on a very small scale as a simple proof of principal study to investigate if the synthesis of mixed peptide sequences may be possible in future.



Scheme 3.5: Synthesis of mixed nucleoside dimers (**56**) and (**57**).

For both mixed nucleoside dimers, (**56**) and (**57**), the conditions used were exactly the same as those described for the synthesis of the thymidine based peptides (Scheme 3.5). Due to the small scale of these reactions, the crude reaction products were directly analysed by electrospray mass spectrometry.

The coupling of the 5′-protected thymidine derived amino acid (**14**) to 5-methyl-2′-deoxycytidine monomer (**16**) was indeed successful with detection of the mass ion for the mixed dimer (**56**) both in low and high resolution mass spectrometry. Even with the crude material, the mass detected at 1006.3392 was extremely close to the calculated value of 1006.3388.

The coupling of the same thymidine derived amino acid (**14**) to the 2′-deoxyadenosine derived monomer (**36**) was also successful. Again mass spectrometry was carried out on the crude mixture and gave a mass ion peak of 800.2529 which is very close the calculated value of 800.2517.

Due to constraints of time, no further work could be carried out on either the thymidine based peptides or the mixed nucleoside based peptides.

3.4 Conclusions

Solid-phase synthesis of peptides made up of thymidine-derived β-amino acid (**3**) was attempted but abandoned, due to complications both with the synthesis, particularly the sensitivity of the glycosidic bond, and the purification of the peptide. This led us to investigate the coupling of nucleoside-derived β-amino acid monomers in solution.

A synthesis of fully protected peptides made up of thymidine-derived β-amino acid units was carried out successfully in solution. Peptides of two, three and four units in length were synthesised. This work was then expanded further to investigate coupling of a 5′-protected thymidine β-amino acid to a 5-methyl-2′-deoxycytidine β-amino acid and a 3′-azido-2′,3′-deoxyadenosine β-amino acid. This work resulted in the successful generation of two mixed nucleoside dipeptides. This short, proof of principle study indicates that, in the future, it may be possible to synthesise longer peptides of mixed sequences in solution.

Although peptides can clearly be prepared by this procedure there remains room for improvement as the yields of some of the coupling steps are fairly low. The work

could also benefit from full NMR analysis to confirm the structures are indeed correct and to investigate if any secondary structures are being formed in the longer peptides. Removal of the protecting groups from the tri- and tetra-peptides could also allow cyclisation studies to be carried out in the future.

In general, the synthesis of peptides in solution remains a work in progress and future studies should include investigations into more efficient methods of purification and analysis. Purification of the fully protected peptides could easily be carried out by flash chromatography. However, removal of the protecting groups from either end of the peptide caused problems. Other methods of purification for these more polar peptides should be investigated such as preparative scale, reverse phase HPLC. Electrospray mass spectrometry is not suited to these large uncharged peptides and other methods such as MALDI-TOF should be investigated in order to obtain cleaner ionisation.

Chapter 4: Experimental

4.1 Techniques and reagents

NMR spectroscopy

All NMR spectra were recorded on a Bruker AMX-400MHz machine. Chemical shifts are reported in ppm relative to an internal standard of tetramethylsilane (TMS) and coupling constants are reported in Hz. Some chemical shift values for aromatic and fluorenyl protons and carbons have been assigned using spectra simulations from ChemBioDraw Ultra 12.0.

Mass spectrometry

Electrospray spectra were recorded by Mr Allan Mills and Miss Moya McCarron at the University of Liverpool using a Micromass LCT machine via Electrospray Ionisation (EI). All samples were recorded using the ionisation mode specified and were injected using a direct infusion syringe pump. HRMS= High Resolution Mass Spectrometry.

IR spectroscopy

All samples were recorded neat on a Jasco FT-IR 4100 spectrometer equipped with an ATR sample holder.

Analytical thin layer chromatography (tlc)

Tlc was performed on UV_{254} sensitive, silica gel 60 coated, aluminium tlc plates purchased from Merck. The developed chromatographs were visualised with a UV lamp (254 nm). Most tlc plates were stained with *p*-anisaldehyde, with spots turning blue upon heating with the exception of those compounds containing a 5'-DMT group which produced an orange stain. Exceptions to this are tlc plates containing carboxylic acid and amino compounds which were stained with bromocresol green and turned yellow and blue, respectively, on a pale blue background without the application of heat.

Flash column chromatography

The required quantity of silica gel (particle size 40-63 μM, supplied by BDH or Sigma Aldrich) was made into a homogenous slurry with the column eluent and applied to the column over a base layer of sand. The material to be purified was then introduced on to the column in a minimum volume of eluent or as a powder by pre-absorption onto silica. All fractions were collected and analysed by tlc.

Solvents

Unless stated solvents were purchased from Fisher Scientific or BDH. Anhydrous solvents were obtained as follows:

Sigma Aldrich Sure-seal™: *N, N*- dimethylformamide 1L, *N, N*- dimethylacetamide 1L, 1,4-dioxane 1L, pyridine 1L, dichloromethane 1L, methanol 1L, tetrahydrofuran 100mL.

VWR: methanol 250mL, dichloromethane 250mL, tetrahydrofuran 250mL, pyridine 250mL.

Link Technologies: acetonitrile 100mL.

General reagents

Unless stated general reagents were obtained from Sigma Aldrich, Fluka or BDH and used as supplied. Nucleosides were obtained from ChemGenes USA and were co-evaporated with the relevant reaction solvent before use. Hydrogen sulfide was obtained in a cylinder from Sigma Aldrich.

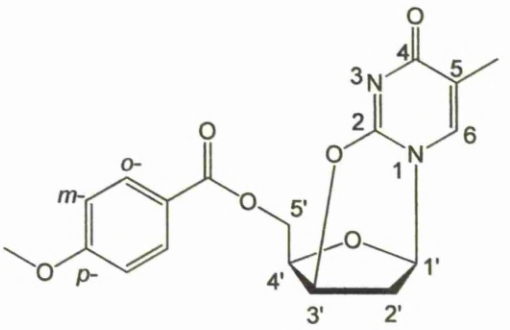
p-Anisaldehyde dip was prepared as follows: *p*-Anisaldehyde (6 mL) was mixed with sulphuric acid (8 mL), acetic acid (2.4 mL) and ethanol (218 mL) and stored at room temperature in darkness.

Bromocresol green dip prepared as follows: Bromocresol green (0.10 g) was dissolved in ethanol (250 mL) and 6M sodium hydroxide (3 mL) was added dropwise. This dip was stored at room temperature.

Lithium azide prepared as follows: Lithium sulfate (10.0 g, 90.9 mmol) and sodium azide (11.8 g, 181.9 mmol) were dissolved in water (40 mL) and warmed to 40°C. Ethanol (60 mL) was then added slowly and the reaction stirred for 10 mins. The solids were removed by filteration under vacuum and the filtrate evaporated to dryness. This afforded lithium azide as a white powdery solid.

4.2 Synthesis of nucleoside monomers

Preparation of 2′, 3′-Anhydro-5′-(4-*O*-methoxybenzoyl)thymidine (**8**)[156]



Thymidine (2.00 g, 8.25 mmol) and PPh_3 (3.20 g, 12.4 mmol) were dissolved in dry DMF (20 mL) and allowed to stir. To this was added *p*-methoxybenzoic acid (1.90 g, 12.4 mmol) in dry DMF (5 mL). The flask was then cooled to 10 °C in an ice/water bath and DIAD (2.50 mL, 12.4 mmol) was added dropwise over 10 minutes, the reaction was then warmed to ambient temperature and allowed to stir for 15 minutes. A further aliquot of PPh_3 (3.20 g, 12.37 mmol) in dry DMF (5 mL) was then added and again the reaction cooled to 10 °C and DIAD (2.50 mL, 12.4 mmol) was added in the same manner as previous. The reaction mixture was then warmed to ambient temperature and allowed to stir for 2 hours. During this time an off -white precipitate formed and the reaction mixture was poured onto Et_2O (250 mL) and cooled to 4°C overnight. The precipitate was then collected by filteration under vacuum and washed with plenty of Et_2O (500 mL).

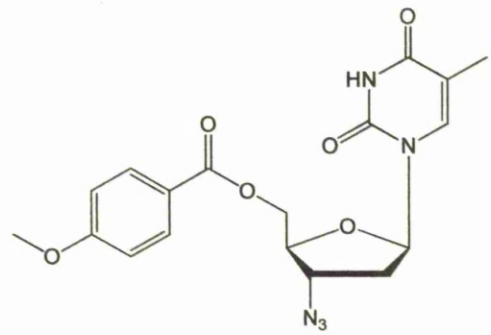
Yield: 2.50 g, 6.98 mmol, 84%

^{1}H NMR (400MHz, DMSO-d_6) δ 1.73 (s, 3H, Thy-CH_3), 2.53 (m, 4H, H2′ + OCH_3), 2.61 (m, 1H, H2′), 4.30 (dd, 1H, *J*= 6.44, 11.81 Hz, H5′), 4.48 (dd, 1H, *J*= 5.21, 11.81 Hz, H5′), 4.57 (m, 1H, H4′), 5.40 (br s, 1H, H3′), 5.88 (apparent d, 1H, *J*= 3.75 Hz, H1′), 7.00 (d, 2H, *J*=6.84 Hz, *m*-ArH), 7.56 (s, 1H, H-6), 7.86 (d, 2H, *J*=6.84 Hz, *o*-ArH).

^{13}C NMR (100MHz, DMSO-d_6) δ 13.3 (Thy-CH_3), 33.1 (C2′), 55.9 (CH_3O), 62.6 (C5′), 77.5 (C3′), 82.4 (C4′), 87.2 (C1′), 114.4 (*m*-ArC), 116.5 (C5), 121.6 (*ipso*-ArC), 131.7 (*o*-ArC), 136.9 (C6), 163.7 (C2), 165.3 (*p*-ArC), 171.2 (C4).

HRMS found m/z (ES^+) 381.1064 ($[M + Na]^+$ 100%); $[C_{18}H_{18}N_2O_6 + Na]^+$ requires 381.1063

Preparation of 3′-Azido-3′-deoxy-5′-*O*-(4-methoxybenzoyl)thymidine (**9**)[156]



2′,3′-Anhydro-5′-(4-*O*-methoxybenzoyl)thymidine (**8**, 2.89 g, 8.06 mmol) and LiN_3 (1.18 g, 24.2 mmol) were dissolved in dry DMF (20 mL) and heated at reflux for 3 hours. The orange reaction mixture was then poured onto 5% aqueous HCl (70 mL). This mixture was then extracted with EtOAc (3 x 100 mL) and the combined organic extracts washed with water (100 mL) followed by brine (100 mL). After drying over $MgSO_4$ all the solvent was removed in vacuo and the crude product was afforded as a yellow gum

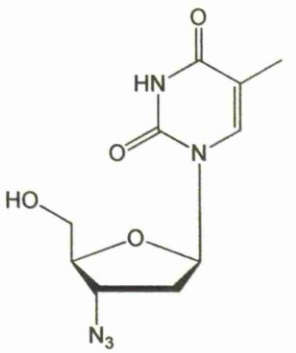
^{1}H NMR (400MHz, $CDCl_3$) δ 1.72 (s, 3H, Thy-CH_3), 2.34 (m, 4H, H2′ + OCH_3), 2.53 (m, 1H, H2′), 4.22 (m, 1H, H4′), 4.32 (m, 1H, H3′), 4.53 (dd, 1H, *J*= 3.80, 12.33 Hz, H5′), 4.65 (dd, 1H, *J*= 3.42, 12.33 Hz, H5′), 6.17 (apparent t, 1H, *J*= 6.46 Hz, H1′), 6.93 (m, 2H, *m*-ArH), 7.20 (s, 1H, H-6), 7.98 (m, 2H, *o*-ArH), 8.16 (br s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 12.6 (Thy-CH_3), 38.3 (C2′), 55.9 (CH_3O), 61.6 (C3′), 63.7 (C5′), 82.6 (C4′), 85.7 (C1’), 99.9 (C5), 114.4 (*m*-ArC), 122.2 (*ipso*-ArC), 132.1 (*o*-ArC), 135.2 (C6), 151.4 (C2), 163.7 (C4), 165.7 (*p*-ArC).

HRMS found m/z (ES^+) 424.1243 ($[M + Na]^+$ 100%); $[C_{18}H_{19}N_5O_6 + Na]^+$ requires 424.1233

IR *v* max/cm^{-1} =1020.16, 1168.65, 1255.43, 1367.28, 1513.85, 1631.48, 1695.12, 2103.96, 2923.56.

Preparation of 3′-Azido-3′-deoxythymidine (**2**, AZT)[156]



3′-Azido-3′-deoxy-5′-*O*-(4-methoxybenzoyl)thymidine (**9**,1.10 g, 2.74 mmol) was dissolved in methanol (15 mL) and NaOMe (590 mg, 11.0 mmol) in water (5 mL) was added. This mixture was then allowed to stir at ambient temperature for 68 hours. Upon addition of a further quantity of water (5 mL) a white solid was seen to precipitate out (*p*-methoxybenzoic acid). The precipitate removed by filteration under vacuum and a further aliquot of water was added (5 mL), any residual white precipitate formed was removed. The aqueous filtrate was then washed with Et_2O (2 x 25 mL) and the two phases were separated. The aqueous layer was then allowed to stir with Amberlite ion exchange resin for 15 minutes. The resin was removed by filteration and all the solvents were removed in vacuo to give a yellow solid. The crude was purified by silica chromatography using a 20% methanol/ethyl acetate eluent system. Fractions containing product were concentrated to afford AZT as an off-white foam.

Yield: 640 mg, 2.39 mmol, 84%

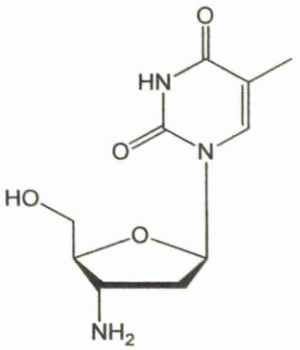
^{1}H NMR (400MHz, MeOD-d_4) δ 1.90 (s, 3H, Thy-CH_3), 2.24 (m, 2H, H2′), 3.76 (dd, 1H, *J*= 3.42, 12.15 Hz, H5′), 3.86 (m, 1H, H5′), 3.93 (m, 1H, H4′), 4.37 (m, 1H, H3′), 6.19 (apparent t, 1H, *J*= 6.35 Hz, H1′), 7.81 (s, 1H, H-6).

^{13}C NMR (100MHz, MeOD-d_4) δ 12.8 (Thy-CH_3), 38.6 (C2′), 61.9 (C5′), 62.8 (C3′), 86.5 (C4′ + C1′), 112.0 (C5), 138.2 (C6), 152.7 (C2), 166.7 (C4).

HRMS found m/z (ES$^+$) 290.0858 ([M + Na]$^+$ 100%); [$C_{10}H_{13}N_5O_4$ + Na]$^+$ requires 290.0865

IR *v* max/ cm^{-1} = 1163.19, 1244.36, 1356.23, 1630.66, 1696.13, 2097.80, 2941.96.

Preparation of 3′-Amino-3′-deoxythymidine (**10**)[156]



3′-Azido-3′-deoxythymidine (**2**, 840 mg, 3.14 mmol) was dissolved in methanol (8 mL) and to this was added a 10% palladium on activated carbon catalyst (20 mg). The reaction vessel was then flushed with nitrogen three times, refilled with H_2 and allowed to stir at ambient temperature and pressure for 4.5 hours. No starting material remained after this time so the reaction was stopped and the Pd catalyst removed by filteration under vacuum through Celite and washed with excess of methanol (100 mL). The solvent was then removed in vacuo to give the product as an off-white foam.

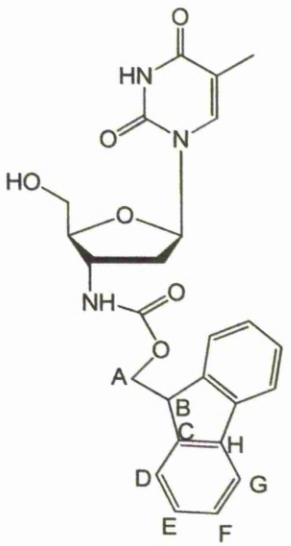
Yield: 555 mg, 2.3 mmol, 74%

^{1}H NMR (400MHz, MeOD-d_4) δ 1.89 (s, 3H, Thy-CH_3), 2.23 (m, 1H, H2′), 2.32 (m, 1H, H2′), 3.33 (m, 1H, H3′), 3.73 (m, 1H, H4′), 3.80-3.89 (m, 2H, H5′), 6.21 (dd, 1H, *J*= 4.84, 6.93Hz, H1′), 7.87 (s, 1H, H-6).

^{13}C NMR (100MHz, MeOD-d_4) δ 12.8 (Thy-CH_3), 42.1 (C2′), 52.0 (C3′), 62.5 (C5′), 86.3 (C1′), 89.0 (C4′), 111.6 (C5), 138.6 (C6), 152.9 (C2), 166.7 (C4).

HRMS found m/z (CI + NH_3) 242.1145 ($[M+H]^+$ 100%); $[C_{10}H_{15}N_3O_4 + H]^+$ requires 242.1140.

Preparation of 3′-*N*-(9-Fluoromethoxycarbonyl)-3′-deoxythymidine (**11**)[156]



3'-Amino-3'-deoxythymidine (**10**, 555 mg, 2.30 mmol) was dissolved in 1,4-dioxane:water (2:1, 9 mL), K_2CO_3 (635 mg, 4.60 mmol) was added and the reaction allowed to stir at ambient temperature for 30 minutes. This was then cooled to 0°C in an ice bath and Fmoc chloride (722 mg, 2.80 mmol) added to the reaction. The mixture was then removed from the ice bath and allowed to stir for 40 minutes at ambient temperature after which the solution appeared a thick semi opaque white colour. Water (10 mL) was added to the reaction and a white precipitate formed which subsequently turned into an off-white sticky solid. Therefore the solvents were decanted off and the solid redissolved in 1% methanol/ethyl acetate (10 mL). All the solvents were removed in vacuo to give an off white solid which was then redissolved in methanol (5 mL) and the remaining insoluble impurities were filtered off. The filtrate was concentated in vacuo and the crude product purified by silica chromatography using a 1% methanol/ethyl acetate eluent system. The fractions containing product were combined to afford a white crystalline solid.

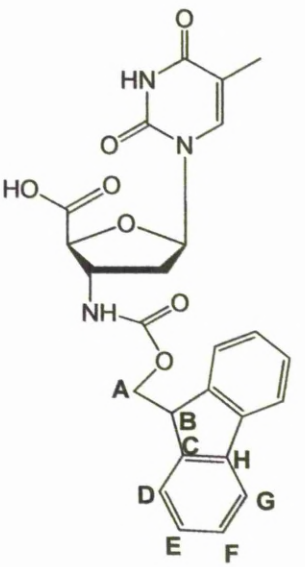
Yield: 480 mg, 1.03 mmol, 45%

^{1}H NMR (400MHz, MeOD-d_4) δ 1.86 (s, 3H, Thy-CH_3), 2.29 (m, 2H, H2′), 3.70 (m, 1H, H4′), 3.85 (m, 2H, H5′), 4.22 (t, 1H, *J*=6.46 Hz, Fmoc-H**B**), 4.28 (m, 1H, H3′), 4.43 (m, 2H, Fmoc-H**A**), 6.20 (m, 1H, H1′), 7.43 (m, 4H, Fmoc-H**E,F**), 7.65 (m, 2H, Fmoc-H**D**), 7.81 (m, 2H, Fmoc-H**G**), 7.91 (s, 1H, H-6).

^{13}C NMR (100MHz, MeOD-d_4) δ 12.8 (Thy-CH_3), 39.3 (C2′), 49.2 (Fmoc C**B**), 52.0 (C3′), 62.7 (C5′), 68.0 (Fmoc C**A**), 86.2 (C1′ + C4′), 112.2 (C5), 121.3, 126.5, 128.5, 129.2 (Fmoc-C**D,E,F,G**), 138.5 (C6), 152.5 (C2), 166.3 (C4).

HRMS found m/z (ES^+) 486.1651 ([M + Na]$^+$ 100%); [$C_{25}H_{25}N_3O_6$ + Na]$^+$ requires 486.1641

Preparation of 3′-*N*-(9-Fluoromethoxycarbonyl)-3′-azido-3′-deoxythymidyl-5′-carboxylic acid (**3**)[156]



3′-*N*-(9-Fluoromethoxycarbonyl)-3′-amino -3′-deoxythymidine (**11**, 480 mg, 1.03 mmol) was suspended in MeCN/water (1:1, 8 mL) and TEMPO (31 mg, 0.20 mmol) and BAIB (728 mg, 2.26 mmol) were added. This caused the suspension to turn bright orange in colour. The reaction was allowed to stir at ambient temperature for 14 hours and after this time the reaction had separated into two distinct layers and a pale brown precipitate was seen to have formed. The precipitate was collected by filteration under vacuum and washed with an excess of Et_2O (100 mL).

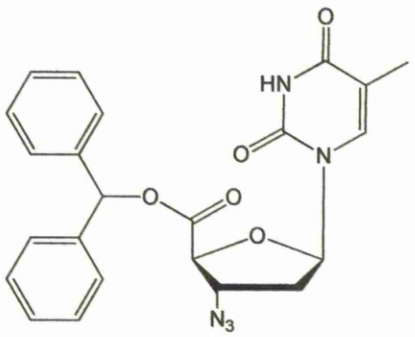
Yield: 200 mg, 0.42 mmol, 40%

^{1}H NMR (400MHz, DMSO-d_6) δ 1.89 (s, 3H, Thy-CH_3), 2.30 (br s, 2H, H2′), 4.22 (m, 1H, Fmoc-H**B**), 4.30-4.50 (4H, m, H4′ + H3′ + Fmoc-H**A**), 6.34 (m, 1H, H1′), 7.35 (m, 2H, Fmoc-H**E**), 7.46 (m, 2H, Fmoc-H**F**), 7.70 (d, 2H, *J*=8.35 Hz, Fmoc-H**D**), 7.80 (d, 2H, *J*= 7.22 Hz, Fmoc –H**G**), 8.19 (s, 1H, 3′-NH), 11.38 (br s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 12.7 (Thy-CH_3), 36.6 (C2′), 47.0 (Fmoc- C**B**), 54.8 (C3′), 65.2 (Fmoc C**A**), 81.9 (C4′), 85.4 (C1′),109.7 (C5), 120.4, 125.5, 127.4, 128.0 (Fmoc-C**D,E,F,G**), 136.7 (C6), 141.1 (Fmoc –C**H**), 144.2 (Fmoc-C**C**), 150.8 (C2), 164.0 (C4), 172.5 (C5′).

HRMS found m/z (ES^+) 500.1429 ([M +Na]$^+$ 100%); [$C_{25}H_{23}N_3O_7$ +Na]$^+$ requires 500.1434.

Preparation of 3′-Azido-3′-deoxythymidine-5′-benzhydryl ester (**13**)



3′-Azido-3′-deoxythymidine (**2**, 2.55 g, 9.54 mmol), TEMPO (297 mg, 1.90 mmol) and BAIB (6.89 g, 21.0 mmol) were suspended in an acetonitrile: water (1:1, 20 mL) mix and allowed to stir at ambient temperature for 1.5 hours. The solvents were removed by in vacuo and residual water removed by co-evaporation with acetone, this afforded the intermediate as a white sticky solid. 3′-Azido-3′-deoxythymidine-5′-carboxylic acid (**12**,1.00 g, 3.56 mmol) and benzophenone hydrazone (965 mg, 5.70 mmol) were added to a flask suspended in an ice bath and DCM (20 mL) was added slowly with stirring. 1% Iodine in DCM (0.2 mL) was then added to the reaction and a mixture of Oxone™ (3.28 g, 5.34 mmol) and wet alumina (3.8 g) was added in portions over 30 mins. The reaction was then allowed to stir in an ice bath over 2 hours and all the solids removed by filteration and washed well with DCM (100 mL). The filtrate was concentrated in vacuo and the crude residue purified by silica chromatography using a 0-2% methanol/DCM eluent system. The desired product was afforded as a yellow foam.

Yield: 1.00 g, 2.09 mmol, 59%

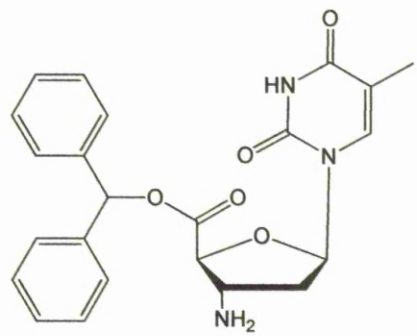
^1H NMR (400MHz, $CDCl_3$) δ 1.84 (s, 3H, Thy-CH_3), 2.13 (m, 1H, H2′), 2.45 (m, 1H, H2′), 4.36 (apparent dt, 1H, *J*= 2.21, 5.99 Hz, H3′), 4.46 (d, 1H, *J*= 2.21 Hz, H4′),

6.37 (dd, 1H, *J*= 5.50, 8.18 Hz, H1'), 7.22-7.40 (m, 11H, Ar-H), 7.80 (d, 1H, *J*=1.24 Hz, H6).

^{13}C NMR (100MHz, $CDCl_3$) δ 12.9 (Thy-CH_3). 37.0 (C2'), 64.2 (C3'), 79.2 (C4'), 86.9 (C1'), 111.9 (C5), [128.8, 128.9, 129.0, 129.2, 129.3 (Ar-C)], 136.1 (C6), 150.6 (C2), 163.93 (C4), 169.6 (C5')

HRMS found m/z (ES^+) 470.1430 ([M + Na]$^+$ 100%); [$C_{23}H_{21}N_5O_5$ + Na]$^+$ requires 470.1440

Preparation of 3'-Amino-3'-deoxythymidine-5'-benzhydryl ester (**14**)



3'-Azido-3'-deoxythymidine-5'-benzhydryl ester (**13**, 1.00 g, 2.09 mmol) was dissolved in a 15 % triethylamine/pyridine solution (20 mL) and cooled to 0°C. Hydrogen sulfide was then bubbled through for 5 minutes after this time the reaction vessel was sealed and allowed to stir for 30 mins at 0°C. The solvents were removed in vacuo and the crude residue purified by silica chromatography using a 1-5% methanol/DCM eluent system. The desired product was afforded as an off-white foam.

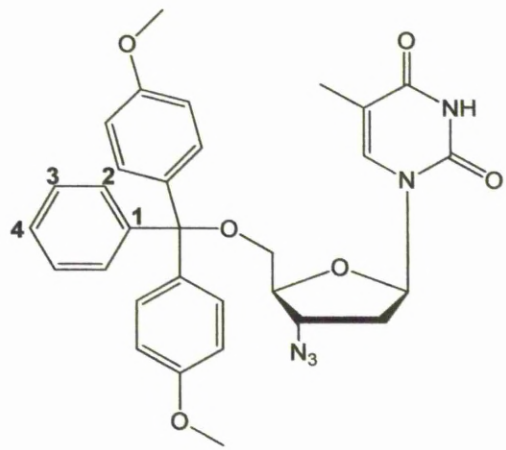
Yield: 780 mg, 1.85 mmol, 88%

^{1}H NMR (400MHz, $CDCl_3$) δ 1.81 (s, 3H, Thy-CH_3), 2.13 (m, 1H, H2'), 2.23 (m, 1H, H2'), 3.76 (apparent dt, 1H, *J*= 3.79, 6.03 Hz, H3'), 4.40 (d, 1H, *J*= 3.79 Hz, H4'), 6.43 (apparent t, 1H, *J*= 6.34 Hz, H1'), 7.00 (s, 1H, Ph_2CH), 7.29-7.39 (m, 10H, Ar-H), 7.94 (d, 1H, *J*= 1.20 Hz, H6)

^{13}C NMR (100MHz, $CDCl_3$) δ 12.9 (Thy-CH_3), 40.4 (C2'), 55.9 (C3'), 78.6 (C4'), 86.8 (C1'), 111.4 (C5), [127.5, 128.9, 129.1, 129.1, 129.2 (Ar-C)], 136.4 (C6), 150.8 (C2), 164.18 (C4), 170.9 (C5')

HRMS found m/z (ES^+) 444.1518 ([M + Na]$^+$ 100%); [$C_{23}H_{23}N_3O_5$ + Na]$^+$ requires 444.1535.

Preparation of 3′-Azido-3′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)thymidine (**17**)[192]



3′-Azido-3′-deoxythymidine (**2**, 3.55 g, 13.2 mmol), Dimethoxytrityl chloride (5.60 g, 15.9 mmol) and DMAP (162 mg, 1.33 mmol) were dissolved in anhydrous pyridine (50 mL) and anhydrous triethylamine (9.24 mL) added. The reaction was then allowed to stir at room temperature for 5 hours. After this time all solvents were removed in vacuo and the residue redissolved in DCM (50 mL), the organic layer was then washed successively with saturated $NaHCO_3$ solution (100 mL) and water (2 x 100 mL). The organics were then dried over $MgSO_4$ and all solvents removed in vacuo. The crude was purified by silica chromatography using a 0-5% methanol/DCM eluent system and fractions containing product were concentrated to afford an off-white foam.

Yield: 6.64 g, 11.7 mmol, 87%

^{1}H NMR (400MHz, MeOH-d_4) δ 1.46 (s, 3H, Thy-CH_3), 2.42 (m, 1H, H2′), 2.51 (m, 1H, H2′) 3.35 (m, 2H, H5′), 3.76 (s, 6H, DMT-OCH_3), 3.94 (m,1H, H4′), 4.46 (m, 1H, H3′), 6.17 (apparent t, 1H, *J*= 6.33 Hz, H1′), 6.86 (m, 4H, DMT-H, *o*-OCH_3), 7.20-7.46 (m, 9H, DMT-H, *m*-OCH_3, **2**, **3**, **4**), 7.65.(s, 1H, H-6).

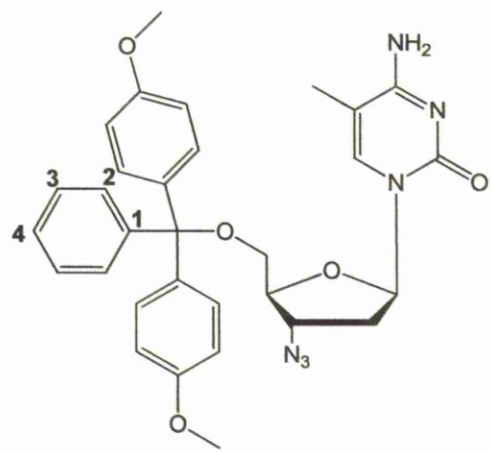
^{13}C NMR (100MHz, MeOD-d_4) δ 12.6 (Thy-CH_3), 38.8 (C2′), 56.1 (DMT-OCH_3 x2), 62.3 (C3′), 64.6 (C5′), 85.3 (C4′), 86.4 (C1′), 88.6 (DMT-qC), 112.1 (C5), 114.7 (DMT-C, *o*-CH_3 x4), 125.9 (DMT-C**4**), 128.5 (DMT-C, *m*-OCH_3 x4), 129.4 (DMT-C**2**),

129.8 (DMT-C**3**), 137.1 (C6), 138.0 (DMT-C, *p*-OCH_3 x2), 146.3 (DMT-C**1**), 150.5 (C2), 160.8 (DMT–C, *ipso*-OCH_3), 166.7 (C4).

HRMS found m/z (ES^+) 592.2169 ($[M + Na]^+$ 100%); $[C_{31}H_{31}N_5O_6 + Na]^+$ requires 592.2172.

IR *v* max/cm^{-1} = 1174.44, 1247.72, 1461.78, 1506.13, 1606.41, 1683.55, 2100.10, 2937.06, 3037.34.

Preparation of 3′-Azido-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**18**)[187,192]



Method 1

Triazole (2.66 g, 38.6 mmol) and phosphoryl chloride (0.78 mL, 8.42 mmol) were suspended in anhydrous MeCN (30 mL) and triethylamine (4.89 mL, 35.1 mmol) was added dropwise at 0˚C. 3′-Azido- 3′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)thymidine (**17**, 2.00 g, 3.51 mmol) in anhydrous MeCN (20 mL) was then added dropwise to the mixture, warmed to room temperature and allowed to stir for 3 hours. The reaction was then cooled to 0˚C, triethylamine (2 mL) and water (0.8 mL) added and the reaction allowed to stir for a further 10 mins. All the solvents were then removed in vacuo and the residue redissolved in EtOAc (50 mL). The organics were washed successively with water (2 x 50 mL) and brine (2 x 50 mL) then dried over $MgSO_4$. The resulting solution was evaporated to dryness then redissolved immediately in 1,4-dioxane (20 mL) and concentrated ammonia (7mL) added, the reaction was then allowed to stir for 3 hours at room temperature. The solvents were removed in

vacuo and the crude was purified by silica chromatography using a 2-5% methanol/DCM eluent system. The product was afforded as a pale brown foam.

Yield: 1.00 g, 1.76 mmol, 50%

Method 2

3′-Azido- 3′-deoxy-5′-*O*-(4,4′-dimethoxytrityl)thymidine (**17**, 2.00 g, 3.51 mmol), 1,2,4-Triazole (727 mg, 10.5 mmol) and 2-chlorophenylphosphorodichloridate (0.86 mL, 5.27 mmol) were dissolved in dry pyridine (20 mL) and allowed to stir under nitrogen at room temperature for 3 days. All pyridine was then removed in vacuo. The residues were then redissolved in 1,4-dioxane (20 mL) and concentrated ammonia was added (7 mL) this was then allowed to stir at ambient temperature for 2 hours. All solvents were then removed in vacuo and the crude residues purified by silica chromatography using a 0-5% methanol/DCM eluent system. The product was afforded as a pale brown foam.

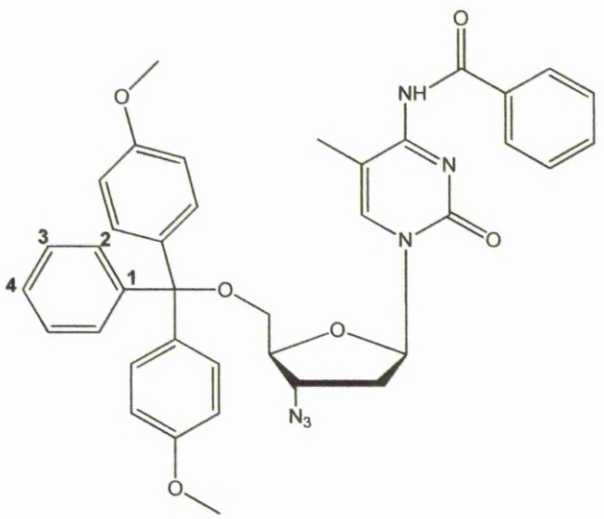
Yield: 1.76 g, 3.10 mmol, 88%

^{1}H NMR (400MHz, $CDCl_3$) δ 1.53 (s, 1H, 5-Me), 2.45 (m, 1H, H2′), 2.60 (m, 1H, H2′), 3.34 (dd, 1H, *J*= 2.89, 10.90 Hz, H5′), 3.60 (dd, 1H, *J*= 2.83, 10.90 Hz, H5′), 3.8 (s, 6H, DMT-OCH_3), 4.02 (m, 1H, H4′), 4.32 (m, 1H, H3′), 6.29 (apparent t, 1H, *J*=5.90 Hz, H1′), 6.88 (m, 4H, DMT-H, *o*-OCH_3), 7.25-7.43 (m, 9H, DMT-H, *m*-OCH_3, **2**, **3**, **4**), 7.78 (s, 1H, H-6), 8.21 (s, 2H, NH_2).

^{13}C NMR (400MHz, $CDCl_3$) δ 13.1 (5-CH_3), 39.1 (C2′), 55.6 (C3′ + DMT-OCH_3 x2), 60.2 (C5′), 84.1 (C4′), 85.9 (C1′), 87.3 (DMT -qC), 102.9 (C5), 113.7 (DMT-C, *o*-OCH_3 x2), 113.7 (DMT-C, *o*-OCH_3 x2), 127.6 (DMT-C**4**), 128.4 (DMT-C, *m*-OCH_3 x4), 130.4 (DMT –C**3** x2 + DMT-C**2** x2), 135.6 (DMT –C, *p*-OCH_3), 135.7 (DMT-C, *p*-OCH_3), 138.2 (C6), 144.6 (DMT-C**1**), 156.5 (C2), 159.1 (DMT-C, *ipso*-OCH_3 x2), 166.2 (C4).

HRMS found m/z (ES^+) 591.2339 ($[M + Na]^+$ 100%); $[C_{31}H_{32}N_6O_5 + Na]^+$ requires 591.2332.

Preparation of *N*-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**15**)[192]



3′-Azido-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**18**, 1.76 g, 3.10 mmol) was dissolved in anhydrous pyridine (15 mL) and cooled to 0˚C. Benzoyl chloride (0.54 mL, 4.65 mmol) was then added dropwise and the reaction allowed to stir at room temperature overnight. The reaction was again cooled to 0˚C and water (0.5 mL) added. After stirring for 5 mins, ammonia (0.5 mL) was added and the reaction mixture allowed to stir for a further 30 mins. All solvents were then removed in vacuo and the residue redissolved in EtOAc (50 mL). The organics were washed successivley with saturated $NaHCO_3$ (2 x 100 mL), water (2 x 100 mL) and brine (2 x 100 mL). The organic layer was dried over $MgSO_4$ and all solvents removed. The crude was then purified by silica chromatography using a 10% methanol/DCM eluent system. The product was afforded as a white foam.

Yield: 1.46 g, 2.17 mmol, 70%

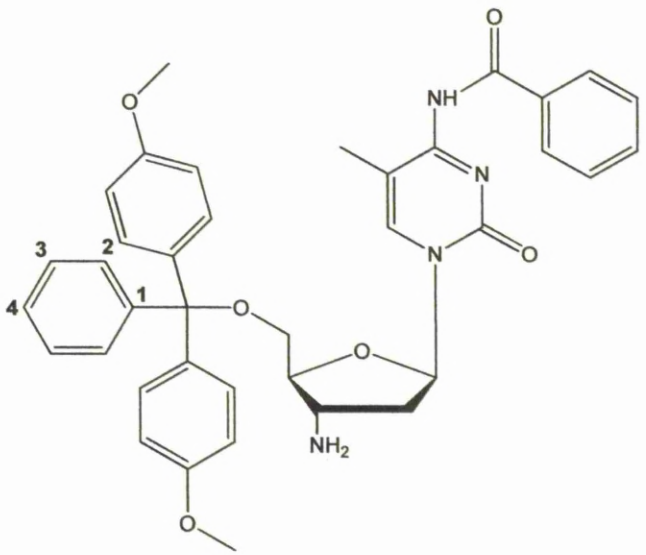
^{1}H NMR (400MHz, $CDCl_3$) δ 1.67 (s, 3H, 5-Me), 2.47 (m, 1H, H2′), 2.54 (m, 1H, H2′), 3.34 (dd, 1H, *J*= 4.08, 10.95 Hz, H5′), 3.60 (dd, 1H, *J*=2.74, 10.95 Hz, H5′), 3.80 (s, 6H, DMT-OCH_3), 4.01 (m,1H, H4′), 4.36 (m, 1H, H3′), 6.28 (apparent t, 1H, *J*= 6.22 Hz, H1′), 6.86 (m, 4H, DMT-H, *o*-OCH_3), 7.24-7.55 (m, 14H, DMT-H, *m*-OCH_3, 1, 2, 3, Bz-H), 7.80 (s, 1H, H-6)

^{13}C NMR(400MHz, $CDCl_3$) δ 13.4 (5-CH_3), 38.7 (C2′), 55.6 (DMT-OCH_3 x2), 60.7 (C5′), 67.2 (C3′), 84.2 (C4′), 85.4 (C1′), 87.5 (DMT-qC), 112.6 (C5), 113.7 (DMT-C, *o*-OCH_3 x2), 113.7 (DMT-C, *o*-CH_3 x2), 127.6 (Bz *o*-C x2 + DMT-C**4**), 128.4 (DMT-C, *m*-OCH_3 x2), 128.5 (DMT-C, *m*-OCH_3 x2), 130.3 (Bz *m*-C), 130.4 (Bz *m*-C),

130.4 (DMT-C**2** x2), 132.8 (DMT-C**3** x2, Bz *p*-C), 135.5 (DMT –C, *p*-OCH_3), 135.5 (DMT-C, *p*-OCH_3), 136.8 (C6), 137.5 (C2), 144.5 (DMT-C**1**), 159.2 (DMT-C, *ipso*-OCH_3 x2), 160.0 (C4).

HRMS found m/z (ES^+) 695.2585 ($[M + Na]^+$ 100%); $[C_{38}H_{36}N_6O_6 + Na]^+$ requires 695.2594.

Preparation of *N*-4-Benzoyl-3′-amino-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**19**)



N-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**15**, 500 mg, 0.80 mmol) was dissolved in anhydrous 15% triethylamine/pyridine (2 mL) and cooled to 0°C in an ice bath. Hydrogen sulfide gas was then bubbled through for 15 mins after which time the reaction vessel was sealed and allowed to stir at ambient temperature for a further 30 mins. All the reaction solvents were removed in vacuo and the crude residue was purified by silica chromatography using a 0-3% methanol/DCM eluent system. The product was afforded as a yellow foam.

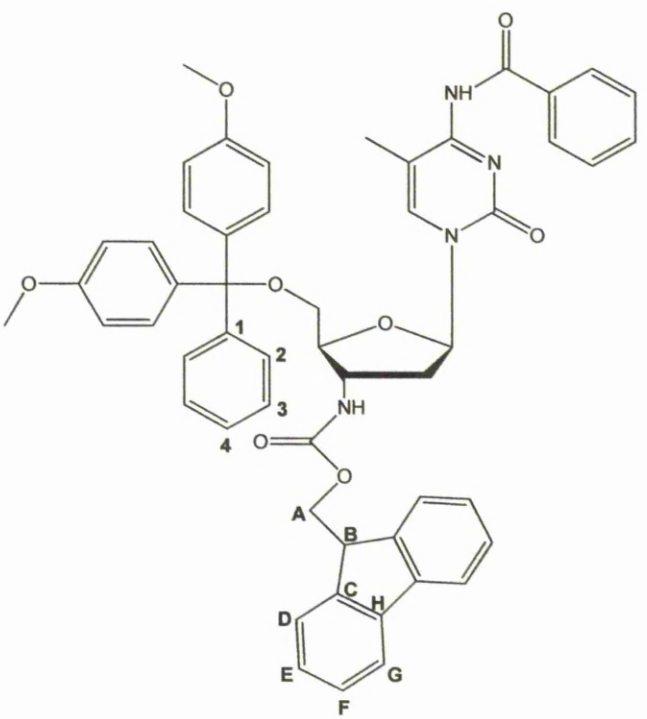
Yield: 363 mg, 0.57 mmol, 78%

^1H NMR (400MHz, $CDCl_3$) δ 1.73 (s, 3H, 5-Me), 2.29 (m, 1H, H2′), 2.45 (m, 1H, H2′), 3.37 (m, 1H, H5′), 3.57 (m, 1H, H5′), 3.77 (m, 2H, H4′ + H3′), 3.82 (s, 6H, DMT-OCH_3), 6.28 (apparent t, 1H, *J*= 6.56 Hz, H1′), 6.88 (m, 4H, DMT-H, *o*-OCH_3), 7.31-7.54 (m, 14H, DMT-H, *m*-OCH_3, **2**, **3**, **4**, Bz-H), 7.84 (s, 1H, H-6), 8.29 (br s, 2H, NH_2).

^{13}C NMR (400MHz, $CDCl_3$) δ 13.5 (5-CH_3), 42.6 (C2′), 51.5 (C3′), 55.6 (DMT-OCH_3 x2), 62.9 (C5′), 85.6 (C4′), 87.1 (C1′), 87.3 (DMT-qC), 113.7 (C5 + DMT-C, *o*-OCH_3 x2), 113.7 (DMT-C, *o*-CH_3 x2), 127.5 (Bz *o*-C x2 + DMT-C**4**), 128.4 (DMT-C, *m*-OCH_3 x2), 128.5 (DMT-C, *m*-OCH_3 x2), 128.5 (Bz *m*-C), 130.2 (Bz *m*-C), 130.5 (DMT-C**2** x2), 132.7 (DMT-C**3** x2, Bz *p*-C), 135.8 (DMT–C, *p*-OCH_3), 136.8 (C6), 137.6 (C2), 144.8 (DMT-C**4**), 159.1 (DMT-C, *ipso*-OCH_3 x2).

HRMS found m/z (ES^+) 669.2667 ($[M + Na]^+$ 100%); $[C_{38}H_{38}N_4O_6 + Na]^+$ requires 669.2689.

Preparation of *N*-4-Benzoyl-3′-(Fmoc)amino-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**20**)



N-4-Benzoyl-3′-amino-2′, 3′-dideoxy-5′-*O*-(4,4′-dimethoxytrityl)-5-methylcytosine (**19**, 100 mg, 0.16 mmol) and potassium carbonate (43 mg, 0.32 mmol) were dissolved in a dioxane/water (2:1, 3 mL) mix and allowed to stir at ambient temperature for 30 mins. The reaction was then cooled to 0°C in an ice bath and Fmoc chloride (49 mg, 0.19 mmol) was added. The reaction was allowed to stir in the ice bath for 2 hours after which time water (1 mL) was added and the reaction stirred for a further 5 mins. All the reaction solvents were removed in vacuo and the crude purified by

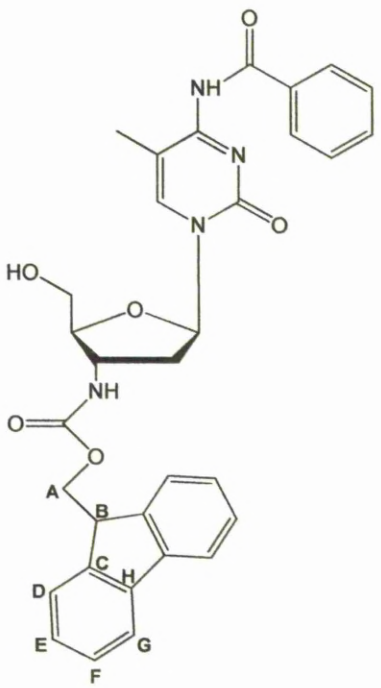
silica chromatography using a 0-2% methanol/DCM eluent system. The product was afforded as an off white solid.

Yield: 140 mg, 0.16 mmol, 99%

^{1}H NMR (400MHz, $CDCl_3$) δ 1.58 (s, 3H, 5-Me), 2.44 (m, 2H, H2′), 3.40 (m, 1H, H5′), 3.52 (m, 1H, H5′), 3.75 (s, 6H, DMT-OCH_3), 3.93 (m, 1H, H3′), 4.19 (m, 1H, H4′), 4.50 (m, 2H, Fmoc-H**A**), 4.98 (m, 1H, Fmoc-H**B**), 6.31 (m, 1H, H1′), 6.84 (m, 4H, DMT-H, *o*-OCH_3), 7.20-7.59 (m, 22H, Fmoc-H**D**, **E**, **F**, **G** + Bz-H +DMT-H), 7.78 (s, 1H, H-6), 8.47 (s, 1H, NH).

HRMS found m/z (ES^+) 891.3351 ([M + Na]$^+$ 100%); [$C_{53}H_{48}N_4O_8$ + Na]$^+$ requires 891.3370.

Preparation of *N*-4-Benzoyl-3′-(Fmoc)amino-2′, 3′-dideoxy-5-methylcytosine (**21**)



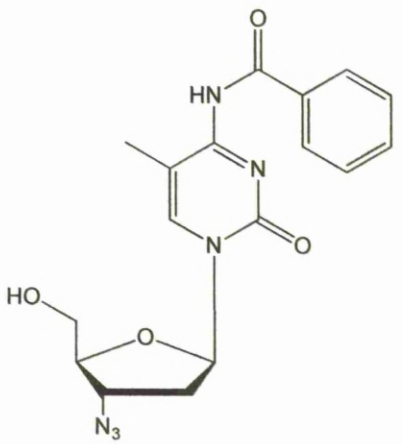
N-4-Benzoyl-3′-(Fmoc)amino-2′, 3′-dideoxy-5′-*O*-(dimethoxytrityl)-5-methylcytosine (**20**, 140 mg, 0.16 mmol) was dissolved in THF (1 mL) and *p*-toluene sulphonic acid (31 mg, 0.18 mmol) in THF (1 mL) was added. The reaction was allowed to stir at ambient temperature for 2 hours and then all reaction solvents were removed in vacuo. The crude residue was purified by silica chromatography using a 0-3% methanol/DCM eluent system. The product was afforded as a white solid.

Yield: 69 mg, 0.12 mmol, 76%

^{1}H NMR (400MHz, DMSO-d_6) δ 2.01 (s, 3H, 5-Me), 2.27 (m, 2H, H2′), 3.59 (m, 1H, H5′), 3.69 (m, 1H, H5′), 3.85 (m, 1H, H3′), 4.12 (m, 3H, Fmoc-H**A** + H4′), 4.35 (m, 1H, Fmoc-H**B**), 6.17 (m, 1H, H1′), 7.32-7.85 (m, 13H, Fmoc-H**D**, **E**, **F**, **G** + Bz-H), 7.90 (s, 1H, H-6), 8.17 (s, 1H, NH).

HRMS found m/z (ES$^+$) 589.2063 ([M + Na]$^+$ 100%); [$C_{32}H_{30}N_4O_6$ + Na]$^+$ requires 589.2063.

Preparation of *N*-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5-methylcytosine (**22**)



N-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5′-*O*-(dimethoxytrityl)-5-methylcytosine (**15**, 500 mg, 0.74 mmol) was dissolved in THF (3 mL) and toluene-4-sulphonic acid (139 mg, 0.81 mmol) was added. This was then allowed to stir at ambient temperature for 5 minutes. All solvents were removed in vacuo and the crude residue purified by silica chromatography using a 0-10% methanol/DCM eluent system. The product was afforded as a white foam.

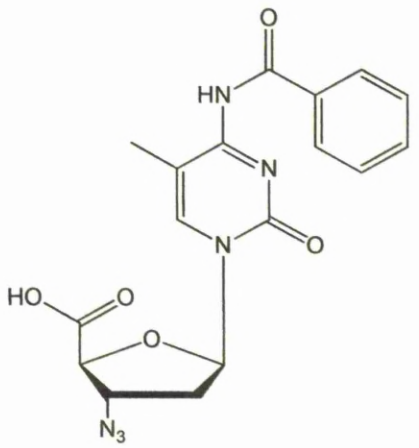
Yield: 220 mg, 0.59 mmol, 81%

^{1}H NMR (400MHz, MeOH-d_4) δ 1.96 (s, 3H, 5-Me), 2.38 (m, 1H, H2′), 2.59 (m, 1H, H2′), 3.76-4.02 (m, 3H, H5′, H4′), 4.37 (m, 1H, H3′), 6.20 (m, 1H, H1′), 7.23 (s, 1H, H-6), 7.70-8.15 (m, 5H, Bz-H)

^{13}C NMR (400MHz, MeOD-d_4) δ 13.6 (5-Me), 39.4 (C2′), 61.9 (C5′), 67.8 (C3′), 86.3 (C4′), 87.5 (C1′), 127.3 (C5), 129.0 (Bz *o*-C x2), 129.9 (Bz *m*-C x2,), 130.1 (Bz *p*-C, Bz *ipso*-C), 130.2 (C6), 130.9 (C2), 142.1 (C4).

HRMS found m/z (ES^+) 393.1296 ($[M + Na]^+$ 100%); $[C_{17}H_{18}N_6O_4 + Na]^+$ requires 393.1287.

Preparation of *N*-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**23**)



N-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5-methylcytosine (**22**, 220 mg, 0.59 mmol) was dissolved in acetonitrile/water (1:1, 3 mL) and cooled to 0˚C in an ice bath. TEMPO (19 mg, 0.12 mmol), BAIB (432 mg, 1.30 mmol) and sodium hydrogen carbonate (100 mg, 1.18 mmol) were then added and the reaction allowed to stir at 0˚C for 2 hours. The reaction was then removed from the ice bath and allowed to stir at room temperature for a further hour, all solvents were then removed in vacuo. The resulting residue was recrystallised from acetone and diethyl ether to give a pale brown solid.

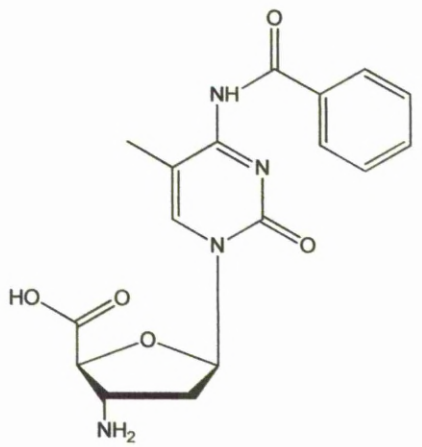
Yield: 188 mg, 0.49 mmol, 83%

^{1}H NMR (400MHz, DMSO-d_6) δ 1.75 (s, 3H, 2-Me), 1.98-2.20 (m, 2H, H2′), 4.02 (m, 1H, H4′), 4.40 (m, 1H, H3′), 6.17 (m, 1H, H1′), 7.10-7.47 (m, 5H, Bz-H), 8.19 (s, 1H, H-6).

^{13}C NMR (400MHz, DMSO-d_6) δ 14.1 (5-Me), 24.7 (C2′), 65.2 (C3′), 85.2 (C4′), 86.6 (C1′), 125.8 (C5), 127.4 (Bz *o*-C x2), 128.4 (Bz *m*-C x2), 128.4 (Bz *p*-C, Bz *ipso*-C), 129.5 (C6), 137.9 (C4), 146.0 (C5′).

HRMS found m/z (ES^-) 383.1110 ($[M - H]^-$ 100%); $[C_{17}H_{15}N_6O_5]^-$ requires 383.1104.

Preparation of *N*-4-Benzoyl-3′-amino-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**24**)



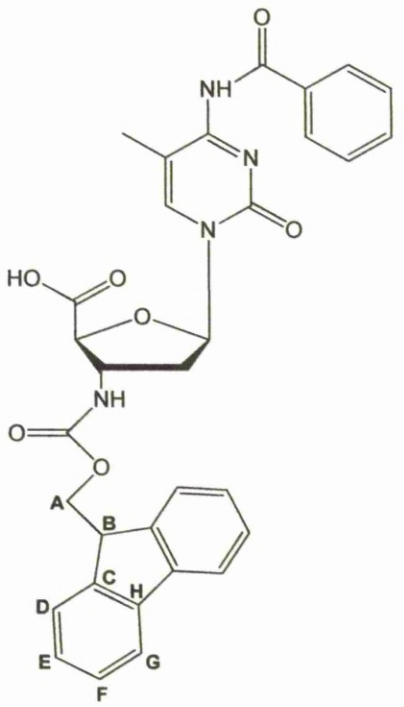
N-4-Benzoyl-3′-azido-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**23**, 180 mg, 0.47 mmol) was suspended in 15% triethylamine/pyridine (3 mL) and cooled to 0˚C. H_2S was then bubbled through the suspension for 20 minutes. After this time the reaction vessel was sealed and allowed to stir for 30 minutes at room temperature. All the solvents were then removed in vacuo and the resulting residue redissolved in water, solid sulphur was seen to precipitate out of solution. The solution was centrifuged and the supernatant fluid evaporated

(Note: This was not analysed by NMR due to residual sulphur present and therefore was used in a partially crude state for the next step. The product identity was confirmed by mass spectrometry)

Yield: 158 mg (Crude mixture, yield not calculated)

HRMS found m/z (ES^-) 357.1214 ($[M-H]^-$ 100%); $[C_{17}H_{17}N_4O_5]^-$ requires 357.1199.

Preparation of *N*-4-Benzoyl-3′-*N*-Fmoc-amino-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**16**)



N-4-Benzoyl-3′-amino-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**24**, 566 mg, 1.58 mmol) and potassium carbonate (436 mg, 3.16 mmol) were dissolved in a dioxane/water (2:1, 3 mL) solvent mixture and allowed to stir at room temperature for 30 mins. The reaction was then cooled to 0°C and Fmoc chloride (408 mg, 1.58 mmol) was added, this was then allowed to stir at the same temperature for 10 hours. Water (10mL) was then added to the reaction and stirred for a further 5 mins. All solvents were then removed in vacuo and the residue redissolved in DCM and purified by silica chromatography using a 0-20% methanol/DCM eluent system. The product was afforded as a pale yellow solid.

Yield: 160 mg, 0.28 mmol, 17%

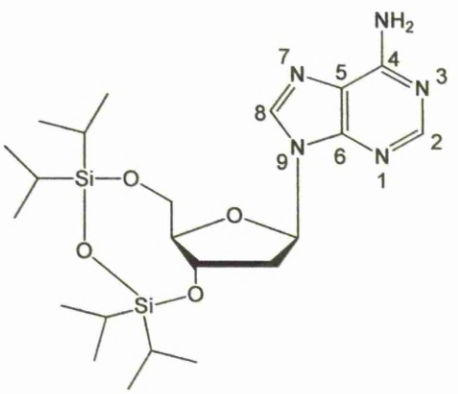
^{1}H NMR (400MHz, DMF-d_7) δ 2.31 (s, 3H, 5-Me), 2.52 (m, 1H, H2′), 2.68 (m, 1H, H2′), 4.49 (m, 1H, H3′), 4.53 (m, 2H, Fmoc-H**A**), 4.70 (m, 2H, H4′ + Fmoc-H**B**), 6.71 (m, 1H, H1′), 7.31 (d, 2H, *J* = 7.59 Hz, Bz *m*-H), 7.50-7.97 (9H, m, Fmoc-H**D**, **E**, **F**, **G** + Bz *p*-H), 8.09 (d, 2H, *J* = 7.59 Hz, Bz *o*-H), 8.47 (s, 1H, H-6).

^{13}C NMR (400MHz, DMF-d_7) δ 13.4 (5-Me), 20.8 (C2′), 47.5 (Fmoc-C**B**), 56.9 (C3′), 66.8 (Fmoc-C**A**), 86.2 (C4′), 87.5 (C1′), 109.2 (C5), 120.4, 120.5, 121.7, 125.8 (Fmoc-C**D**, **E**, **F**, **G**), 126.3 (C5), 127.5 (Bz *o*-C x2), 128.5 (Bz *m*-C x2), 128.8 (Bz *p*-

C), 132.8 (C6), 141.5 (Fmoc-C**H**), 144.7 (Fmoc-C**C**), 156.9 (C4), 182.5 (Bz C=O), 206.4 (C5′ C=O).

HRMS found m/z (ES^-) 579.1885 ($[M-H]^-$ 100%); $[C_{32}H_{27}N_4O_7]^-$ requires 579.1880.

Preparation of 3′,5′-Tetraisopropylsiloxane-2′-deoxyadenosine (**29**)[211]



2′-Deoxyadenosine (10.0 g, 37.1 mmol) was co-evaporated three times with anhydrous pyridine (3 x 150 mL) and then suspended in fresh anhydrous pyridine (80 mL). 1,3-Dichloro-1,1,3,3-tetraisopropyldisiloxane (12.1 mL, 37.9 mmol) was added to the suspension dropwise and the reaction allowed to stir at ambient temperature for 18 hours. All the solvent was removed in vacuo and the residue redissolved in DCM (100 mL). The solution was then washed successively with saturated sodium hydrogen carbonate (200 mL), water (200 mL) and brine (200 mL). The organic layer was dried over $MgSO_4$ and all solvents removed in vacuo. The crude residue was purified by silica chromatography using a 2-4% methanol/DCM eluent system. The desired product was afforded as a white foam.

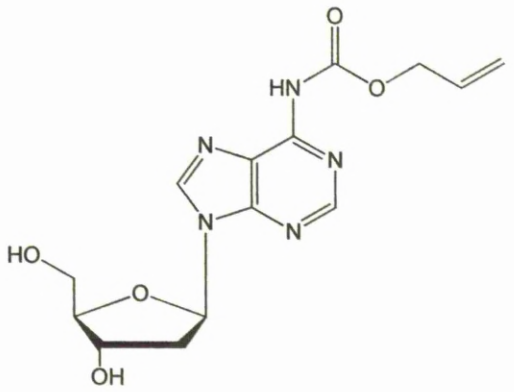
Yield: 16.5 g, 33.3 mmol, 90%

^1H NMR (400MHz, $CDCl_3$) δ 1.30-1.11 (m, 28H, *iso*-Propyl H), 2.67 (m, 2H, H2′), 3.90 (m, 1H, H4′), 4.06 (m, 2H, H5′), 4.95 (apparent-q, 1H, *J*= 7.97 Hz, H3′), 6.30 (dd, 1H, *J*= 2.58 Hz, H1′), 8.04 (s, 1H, H8), 8.32 (s, 1H, H2).

^{13}C NMR (100MHz, $CDCl_3$) δ [12.9, 13.2, 13.5, 13.7, 17.2, 17.3, 17.4, 17.5, 17.7, 17.8, 17.9, (*iso*-propyl C)], 40.4 (C2′), 62.1 (C5′), 70.1 (C3′), 83.5 (C1′), 85.5 (C4′), 120.6 (C5), 139.3 (C8), 149.4 (C2), 153.3 (C6), 155.8 (C4).

HRMS found m/z (ES^+) 516.2435 ($[M + Na]^+$ 100%); $[C_{22}H_{39}N_5O_4Si_2 + Na]^+$ requires 516.2438

Preparation of *N*-6-Allyloxycarbonyl-2'-deoxyadenosine (**25**)



Method A

Ethylthio-tetrazole (7.71 g, 59.3 mmol) was dissolved in a mixture of THF (150 mL) and triethylamine (8.25 mL, 59.3 mmol) and cooled to 0 °C in ice. Allyl chloroformate (6.30 mL, 59.3 mmol) was then added dropwise over a period of 10 mins and the reaction was allowed to stir for 1 hour at 0 °C. The reaction was filtered through celite to remove solids and concentrated to $1/10^{th}$ of its original volume by solvent evaporation. The intermediate formed was then added to a solution of 3',5'-tetraisopropylsiloxane-2'-deoxyadenosine (**29**, 9.75 g, 19.8 mmol) in anhydrous THF (150 mL) and heated to 70°C for 20 hours. After cooling all the solvent was removed in vacuo and the oily residue redissolved in DCM (200 mL). This was then washed successively with saturated sodium hydrogen carbonate (150 mL), water (150 mL) and brine (150 mL) and the organic layer dried over $MgSO_4$. The solvents were removed in vacuo and fresh anhydrous THF (150 mL) was added. TBAF (49.4 mL, 49.4 mmol) was then added cautiously and the reaction allowed to stir at ambient temperature for 48hours. The solvents were then removed in vacuo and the crude residue purified by silica chromatography using a 1-8% methanol/DCM eluent system. The desired product was afforded as a white foam.

Yield: 5.52 g, 16.5 mmol, 83%

Method B[208]

Triethylamine (1.4 mL, 10 mmol) was added dropwise to a 3% solution of tetrazole in acetonitrile (24.0 mL, 10.0 mmol) and the reaction cooled to 0°C in ice. Allyl chloroformate was then added dropwise over 10 mins and the reaction allowed to stir in an ice bath for 30 mins. The reaction was filtered through a pad of celite to

remove solids and concentrated to 1/10th of its original volume in vacuo . The intermediate formed was then added to a solution of 3',5'-tetraisopropylsiloxane-2'-deoxyadenosine (**29**, 1.30 g, 2.63 mmol) in anhydrous THF (35 mL) and heated to 70°C for 2 hours. After cooling all the solvent was removed in vacuo and the oily residue redissolved in ethyl acetate (75 mL). This was then washed with sodium hydrogen carbonate (75 mL) and the organic layer dried over $MgSO_4$. The solvents were removed in vacuo and fresh anhydrous THF (2 0mL) was added. TBAF (5.00 mL, 5.00 mmol) was then added cautiously and the reaction allowed to stir at ambient temperature for 2 hours. The reaction solvents were then removed in vacuo and the crude residue purified by silica chromatography using a 2-8% methanol/DCM eluent system. The desired product was afforded as a white foam.

Yield: 521mg, 1.55mmol, 60%

Method C[211]

2'-Deoxyadenosine (1.00 g, 3.71 mmol) was co-evaporated three times with anhydrous dioxane (3 x 20 mL). It was then suspended in dry dioxane (13 mL) along with hexamethyldisilazane (12.4 mL, 59.4 mmol) and ammonium sulphate (0.20 g, 1.51 mmol) and heated to 115°C. The resulting mixture was allowed to stir for 2 hours. All reaction solvents were then removed in vacuo to give an oily residue. The residue was co-evaporated twice with dry toluene (2 x 50 mL) and then redissolved in anhydrous DCM (40 mL). The solution was cooled in an ice bath and N-methylimidazole (0.83 mL, 10.4 mmol) and allylchloroformate (1.10 mL, 10.4 mmol) were added, the resulting mixture was allowed to stir at ambient temperature for 42 hours. The reaction solvents were removed in vacuo and the residue redissolved in methanol (35 mL) and triethylamine (8 mL) and allowed to stir at ambient temperature for 15 hours. The solvents were removed in vacuo and the crude residue purified by silica chromatography using a 4-10% methanol/ DCM eluent system. The product was afforded as a white foam.

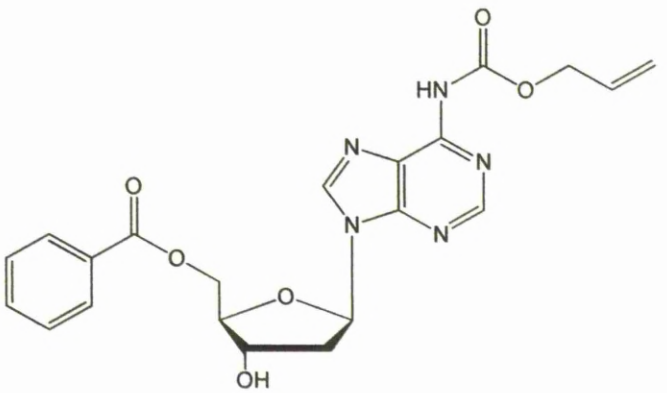
Yield: 150 mg, 0.44 mmol, 11%

^{1}H NMR (400MHz, $CDCl_3$) δ 2.44 (m, 1H, H2'), 2.99 (m, 1H, H2'), 3.84 (dd, 1H, *J*= 2.08, 12.77 Hz, H5'), 3.97 (dd, 1H, *J*= 1.90, 12.77 Hz, H5'), 4.25 (m, 1H, H4'), 4.77 (m, 2H, C$\underline{H}_2$CH=CH_2), 4.82 (m, 1H, H3'), 5.29 (apparent dq, 1H, *J*= 1.12, 10.40 Hz) CH=C$\underline{H}_2$), 5.39 (apparent dq, 1H, *J*= 1.38, 17.18 Hz, CH=C$\underline{H}_2$), 5.99 (m, 1H, C$\underline{H}$=CH_2), 6.44 (dd, 1H, *J*= 5.68, 8.72Hz, H1'), 8.19 (s, 1H, H8), 8.74 (s, 1H, H2), 8.85 (s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 41.2 (C2′), 63.4 (C5′), 67.2 ($\underline{C}H_2CH=CH_2$), 73.1 (C3′), 87.7 (C1′), 89.8 (C4′), 119.6 ($CH=\underline{C}H_2$), 131.9 ($\underline{C}H=CH_2$), 143.0 (C8), 150.3 (C5), 150.5 (C6), 151.2 (C2), 152.5 (C4).

HRMS found m/z (ES^+) 358.1113 ($[M + Na]^+$ 100%); $[C_{14}H_{17}N_5O_5 + Na]^+$ requires 358.1127

Preparation of *N*-6-Allyloxycarbonyl-5′-benzoyl-2′-deoxyadenosine (**30**)



N-6-Allyloxycarbonyl-2′-deoxyadenosine (**25**, 630 mg, 1.88 mmol) was dissolved in anhydrous pyridine (3mL) and cooled in an ice bath. Benzoyl chloride (218 µl, 1.88 mmol) in anhydrous pyridine (2 mL) was added dropwise over a period of 30mins. The reaction mixture was allowed to stir for a further 30 mins. Water (0.3 mL) was added to the reaction and the mixture allowed to stir for 5 mins, all reaction solvents were then removed in vacuo. The green residue was redissolved in DCM (20 mL) and washed with water (20 mL), the organic layer was dried over $MgSO_4$ and all solvents removed in vacuo. The crude residue was purified by silica chromatography using a 0-6% methanol/ DCM eluent system and the desired product was afforded as a white foam.

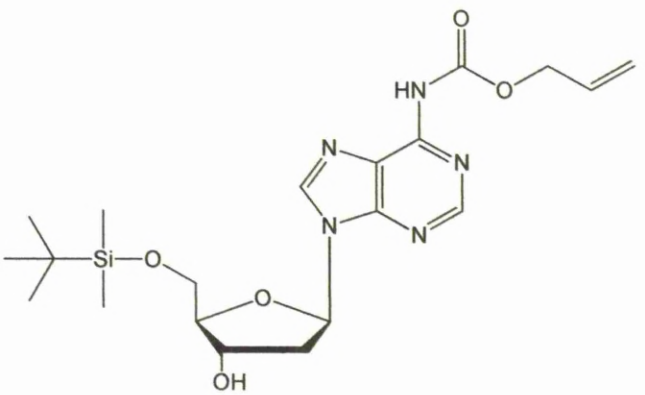
Yield: 370 mg, 0.84 mmol, 45%

^{1}H NMR (400MHz, DMSO–d_6) δ 2.45 (m 1H, H2′), 2.99 (m, 1H, H2′), 4.18 (m, 1H, H4′), 4.44 (dd, 1H, *J*= 5.92, 11.83 Hz, H5′), 4.56 (dd, 1H, *J*= 4.26, 11.83 Hz, H5′), 4.67 (m, 3H, $C\underline{H}_2CH=CH_2$ + H3′), 5.22 (apparent dq, 1H, *J*= 1.47, 10.55 Hz, $CH=C\underline{H}_2$), 5.41 (apparent dq, 1H, *J*= 1.69, 17.24 Hz, $CH=C\underline{H}_2$), 5.60 (d, 1H, *J*= 4.36Hz, 3′-OH), 5.98 (m, 1H, $C\underline{H}=CH_2$), 6.48 (apparent t, 1H, *J*=6.60Hz, H1′), 7.50 (m, 2H, *m*-Ar Bz), 7.64 (m, 1H, *p*-Ar Bz), 7.90 (m, 2H, *o*-Ar Bz), 8.60 (s, 1H, H8), 8.62 (s, 1H, H2), 10.65 (br s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 38.7 (C2′), 64.6 (C5′), 65.6 (C$\underline{H}_2$CH=CH$_2$), 70.7 (C3′), 84.0 (C1′), 84.4 (C4′), 117.9 (CH=C$\underline{H}_2$), 124.2 (C5), 129.0 (*m*-ArC), 129.5 (*o*-ArC), 129.6 (*ipso*-ArC), 133.2 (C$\underline{H}$=CH$_2$), 133.7 (*p*-ArC), 143.3 (C8), 150.0 (C6), 151.7 (C=O Alloc), 152.0 (C2), 152.2 (C4), 165.8 (C=O Bz).

HRMS found m/z (ES$^+$) 462.1392 ([M + Na]$^+$ 100%); [$C_{21}H_{21}N_5O_6$ + Na]$^+$ requires 462.1390

Preparation of *N*-6-Allyloxycarbonyl-5′- TBDMS-2′-deoxyadenosine (**38**)



N-6-Allyloxycarbonyl-2′-deoxyadenosine (**25**, 4.29 g, 12.8 mmol) and TBDMS chloride (2.70 g, 17.9 mmol) were dissolved in anhydrous pyridine (40 mL) and allowed to stir at ambient temperature for 16 hours. The pyridine was removed in vacuo and the resulting residue redissolved in ethyl acetate (100 mL) and washed successively with saturated sodium hydrogen carbonate (150 mL), water (150 mL) and brine (150 mL). The organic layer was dried with $MgSO_4$ and evaporated to dryness. The crude residue was purified by silica chromatography using a 2-6% methanol/DCM eluent system and the desired product was afforded as a white foam.

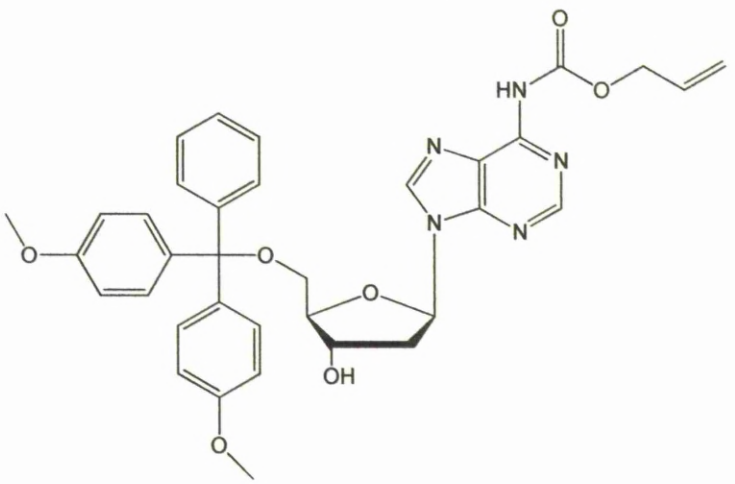
Yield: 3.21g, 7.14mmol, 56%

^{1}H NMR (400MHz, $CDCl_3$) δ 0.01 (s, 6H, 2 x CH_3), 0.80 (s, 9H, *t*-Bu), 2.57 (m, 2H, H2′), 3.78 (dd, 1H, *J*= 3.22, 11.20 Hz, H5′), 3.83 (dd, 1H, *J*= 3.94, 11.20 Hz, H5′), 4.07 (apparent-q, 1H, *J*= 3.44 Hz, H4′), 4.62 (m , 1H, H3′), 4.67 (m, 2H, C$\underline{H}_2$CH=CH$_2$), 5.18 (apparent dq, 1H, *J*= 1.17, 10.40 Hz, CH=C$\underline{H}_2$), 5.30 (apparent dq, 1H, *J*= 1.40, 17.18 Hz, CH=C$\underline{H}_2$), 5.89 (m, 1H, C$\underline{H}$=CH$_2$), 6.47 (apparent t, 1H, *J*= 6.40Hz, H1′), 8.27 (s, 1H, H8), 8.65 (s, 1H, H2), 8.85 (br s, 1H, NH).

^{13}C NMR (400MHz, $CDCl_3$) δ -4.9 (CH_3), 18.7 ($\underline{C}(CH_3)_3$), 26.3 ($C(\underline{C}H_3)_3$), 41.7 (C2′), 63.7 (C5′), 67.0 ($\underline{C}H_2CH=CH_2$), 72.4 (C3′), 85.0 (C1′), 87.7 (C4′), 119.4 ($CH=\underline{C}H_2$), 122.5 (C5), 132.1 ($\underline{C}H=CH_2$), 141.5 (C8), 149.7 (C=O Alloc), 151.0 (C6), 151.2 (C4), 153.1 (C2).

HRMS found m/z (ES^+) 472.1996 ($[M + Na]^+$ 100%); $[C_{20}H_{31}N_5O_5 + Na]^+$ requires 472.1992.

Preparation of *N*-6-Allyloxycarbonyl-5′-(4,4′-dimethoxytrityl)-2′-deoxyadenosine (**40**)



N-6-Allyloxycarbonyl-2′-deoxyadenosine (**25**, 3.50 g, 10.4 mmol), dimethoxytrityl chloride (3.53 g, 10.4 mmol) and dimethoxyaminopyridine (127 mg, 1.04 mmol) were dissolved in anhydrous pyridine (100 mL) and allowed to stir at ambient temperature for 18 hours. All reaction solvent was removed in vacuo and the resulting crude residue purified by silica chromatography using a 1-4% methanol/DCM eluent system containing 1% triethylamine to basify the silica. The desired product was afforded as an off-white foam.

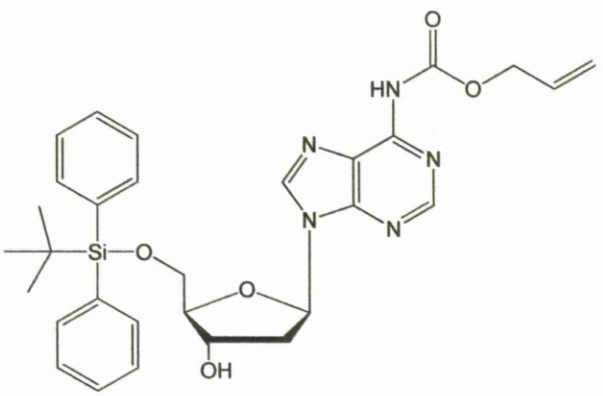
Yield: 2.85 g, 4.47 mmol, 42%

^1H NMR (400MHz, $CDCl_3$) δ 2.57 (m, 1H, H2′), 2.86 (m, 1H, H2′), 3.40 (m, 2H, H5′), 3.77 (s, 6H, 2 x O-CH_3), 4.16 (apparent q, 1H, *J*= 4.32 Hz, H4′), 4.71 (apparent dt, 1H, *J*= 3.93, 6.02 Hz, H3′), 4.76 (m, 2H, $C\underline{H}_2CH=CH_2$), 5.29 (apparent dq, 1H, *J*= 1.21, 10.28 Hz, $CH=C\underline{H}_2$), 5.40 (apparent dq, 1H, *J*= 1.36, 17.10 Hz, $CH=C\underline{H}_2$), 5.99 (m, 1H, $C\underline{H}=CH_2$), 6.47 (apparent t, 1H, *J*= 6.38 Hz, H1′), 6.81 (m, 4H, ArH), 7.13-7.39 (m, 9H, ArH), 8.10 (s, 1H, H8), 8.68 (s, 1H, H2), 8.77 (s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 40.6 (C2′), 55.6 (O-CH_3), 64.0 (C5′), 67.0 ($\underline{C}H_2CH=CH_2$), 72.8 (DMT *q*-C), 85.0 (C1′), 86.5 (C4′), 113.5 (ArC), 113.5 (ArC), 119.4 ($CH=\underline{C}H_2$), 122.8 (C5), [127.4, 128.2, 128.3, 129.2, 129.5, 130.3 (ArC)], 132.1 ($\underline{C}H=CH_2$), 135.9 (ArC), 141.5 (C8), 144.8 (ArC), 149.7 (ArC), 151.1 (C6), 153.1 (C4), 158.9 (C2).

HRMS found m/z (ES^+) 660.2415 ($[M+Na]^+$ 100%); $[C_{35}H_{35}N_5O_7 + Na]^+$ requires 660.2434.

Preparation of *N*-6-Allyloxycarbonyl-5′-TBDPS-2′-deoxyadenosine (**45**)



N-6-Allyloxycarbonyl-2′-deoxyadenosine (**25**, 7.90 g including TBAF salts) and dimethylaminopyridine (575 mg, 4.30mmol) were dissolved in anhydrous pyridine (60 mL) and *tert*-butyldiphenylsilyl chloride (3.06 mL) was added dropwise. The reaction mixture was allowed to stir at ambient temperature under nitrogen for 4 hours after which time no starting material remained. The reaction solvents were removed in vacuo and the residue redissolved in DCM (70 mL). The organic solution was washed successively with a saturated sodium hydrogen carbonate solution (100 mL), water (100 mL) and brine (100 mL). The organic layer was dried over $MgSO_4$ and evaporated to dryness. The residue was purified by silica chromatography using a 0-4% methanol/DCM eluent system. The product was afforded as a white foam.

Yield: 3.10 g, due to starting material impurities accurate yield not recorded.

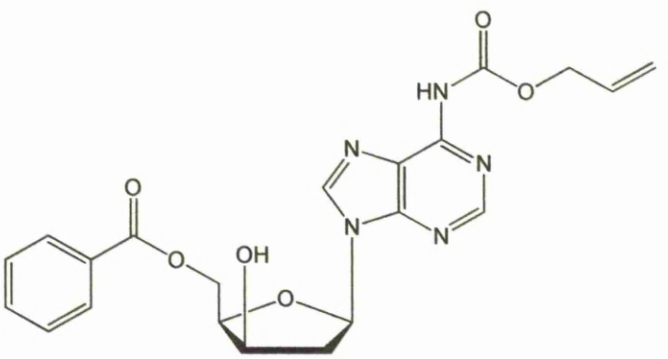
^1H NMR (400MHz, $CDCl_3$) δ 1.04 (s, 9H, *t*-Bu), 2.59 (m, 1H, H2′), 2.76 (m, 1H, H2′), 3.85 (dd, 1H, *J*= 3.86, 11.36 Hz, H5′), 3.94 (dd, 1H, *J*= 4.75, 11.36 Hz, H5′), 4.13 (apparent q, 1H, *J*= 4.04 Hz, H4′), 4.74 (m, 3H, $C\underline{H}_2CH=CH_2$ + H3′), 5.25 (apparent dq, 1H, *J*= 1.20, 10.40 Hz, $CH=C\underline{H}_2$), 5.37 (apparent dq, 1H, *J*= 1.45, 17.40 Hz,

CH=C$\underline{H}_2$), 5.96 (m, 1H, C$\underline{H}$=CH_2), 6.50 (apparent t, 1H, *J*= 6.48 Hz, H1'), 7.37 (m , 6H, ArH), 7.62 (m, 4H, ArH), 8.19 (s, 1H, H8), 8.70 (m, 1H, H2), 8.84 (br s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 19.6 (C($\underline{C}H_3$)$_3$), 27.3 (C($\underline{C}H_3$)$_3$), 40.9 (C2'), 64.2 (C5'), 67.0 ($\underline{C}H_2CH$=CH_2), 72.3 (C3'), 85.0 (C1'), 87.5 (C4'), 119.4 (CH=$\underline{C}H_2$), 122.6 (C5), 132.1 (C$\underline{H}$=CH_2), [133.0, 133.1, 135.8, 135.9, (ArC)], 141.4 (C8), 149.7 (C=O Alloc), 151.1 (C6), 151.2 (C4), 153.1 (C2).

HRMS found m/z (ES$^+$) 596.2286 ([M + Na]$^+$ 100%); [$C_{30}H_{35}N_5O_5Si$ + Na]$^+$ requires 596.2305.

Preparation of *N*-6-Allyloxycarbonyl-5'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**31**) (and *N*-6-allyloxycarbonyl-3'-benzoyl-3'-*xylo*-2'-deoxyadenosine, **32**)



N-6-Allyloxycarbonyl-5'-benzoyl-2'-deoxyadenosine (**30**, 5.15 g, 11.7 mmol) was dissolved in a mixture of anhydrous DCM (50 mL) and Pyridine (2.5 mL) and cooled to -30°C. A solution of triflic anhydride (1.45 mL, 8.64 mmol) in anhydrous DCM (14.5 mL) was then added dropwise over 30 minutes. The reaction was warmed to ambient temperature and water (25 mL) was added. The reaction mixture was allowed to stir for 1 hour before a further portion of water (50 mL) was added and the layers formed were separated. The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude residue was purified using silica chromatography using a 0-2.5% methanol/DCM eluent system to afford both products as yellow foams

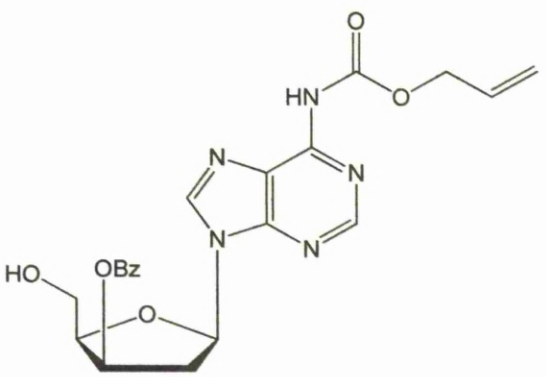
Yield: 583 mg, 1.32 mmol, 11%

^{1}H NMR (400MHz, DMSO-d_6) δ 2.27 (m, 1H, H2'), 2.68 (m, 1H, H2'), 4.17 (m, 1H, H4'), 4.34 (m, 2H, H3' + H5'), 4.47 (m, 3H, $C\underline{H}_2CH$=CH_2 + H5'), 5.50 (apparent dq, 1H, *J*= 1.39, 10.54 Hz, CH=C$\underline{H}_2$), 5.23 (apparent dq, 1H, *J*= 1.62, 17.26 Hz,

CH=C$\underline{H}_2$), 5.70 (d, 1H, *J*= 4.48 Hz, 3′-OH), 5.80 (m, 1H, C$\underline{H}$=CH$_2$), 6.29 (dd, 1H, *J*= 1.68, 8.24 Hz, H1′), 7.33 (m, 2H, *m*-ArH), 7.47 (m, 1H, *p*-ArH), 7.77 (m, 1H, *o*-ArH), 8.46 (s, 1H, H8), 8.51 (s, 1H, H2), 10.49 (br s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 40.9 (C2′), 64.4 (C5′), 65.6 ($\underline{C}$H$_2$CH=CH$_2$), 69.7 (C3′), 82.5 (C4′), 83.4 (C1′), 117.9 (CH=$\underline{C}$H$_2$), 123.8 (C5), 129.1 (*m*-ArC), 129.5 (*o*-ArC), 129.8 (*ipso*-ArC), 133.2 ($\underline{C}$H=CH$_2$), 133.7 (*p*-ArC), 143.2 (C2), 149.9 (C6), 151.6 (C=O Alloc), 151.9 (C8), 152.2 (C4), 166.0 (C=O Bz).

HRMS found m/z (ES$^+$) 462.1393 ([M + Na]$^+$ 100%); [$C_{21}H_{21}N_5O_6$ + Na]$^+$ requires 462.1390.



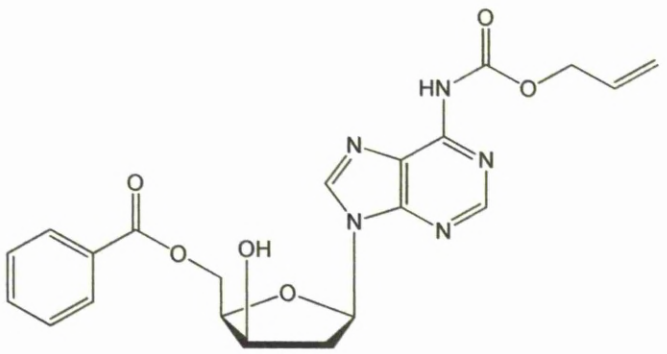
Yield: 1.70 g, 3.87 mmol, 33%

^{1}H NMR (400MHz, DMSO-d_6) δ 2.93 (m, 1H, H2′), 3.06 (m, 1H, H2′), 3.80 (m, 2H, H5′), 4.37 (m, 1H, H4′), 4.70 (m, 2H, C$\underline{H}_2$CH=CH$_2$), 4.99 (t, 1H, *J*= 5.58 Hz, 5′-OH), 5.24 (apparent dq, 1H, *J*= 1.52, 10.52 Hz, CH=C$\underline{H}_2$), 5.42 (apparent dq, 1H, *J*= 1.62, 17.22 Hz, CH=C$\underline{H}_2$), 5.72 (apparent t, 1H, *J*= 4.14 Hz, H3′), 5.97 (m, 1H, C$\underline{H}$=CH$_2$), 6.52 (dd, 1H, *J*= 2.56, 7.52 Hz, H1′), 7.50 (m, 2H, *m*-ArH), 7.65 (m, 1H, *p*-ArH), 7.80 (m, 2H, *o*-ArH), 8.54 (s, 1H, H8), 8.59 (s, 1H, H2), 10.65 (br s, 1H, NH).

^{13}C NMR (400MHz, DMSO-d_6) δ 38.1 (C2′), 59.5 (C5′), 65.5 ($\underline{C}$H$_2$CH=CH$_2$), 73.4 (C3′), 83.8 (C1′ + C4′), 117.9 (CH=$\underline{C}$H$_2$), 124.0 (C5), 129.1 (*m*-ArC), 129.4 (*o*-ArC), 129.8 (*ipso*-ArC), 133.2 ($\underline{C}$H=CH$_2$), 133.7 (*p*-ArC), 142.2 (C2), 149.9 (C6), 151.7 (C=O Alloc), 151.9 (C8), 152.2 (C4), 165.2 (C=O Bz).

HRMS found m/z (ES$^+$) 462.1400 ([M + Na]$^+$ 100%); [$C_{21}H_{21}N_5O_6$ + Na]$^+$ requires 462.1390

Isomerisation of *N*-6-Allyloxycarbonyl-3'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**32**) to *N* -6-allyloxycarbonyl-5'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**31**)

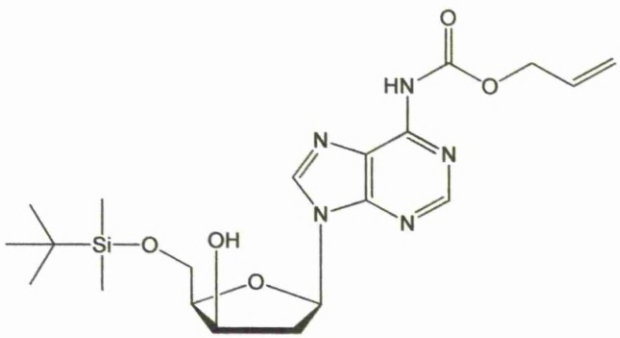


N-6-Allyloxycarbonyl-3'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**32**, 1.70 g, 3.87 mmol) was dissolved in methanol (125 mL) and solid sodium hydrogen carbonate (260 mg, 3.09 mmol) was added. The reaction mixture was allowed to stir at ambient temperature for 2 hours. The reaction solvents were removed in vacuo and the resulting crude residue purified by silica chromatography using a 1-6% methanol/DCM eluent system. The desired product was afforded as an off-white foam.

Yield: 600 mg, 1.36 mmol, 35%

Analysis as for *N* -6-allyloxycarbonyl-5'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**31**).

Preparation of *N*-6-Allyloxycarbonyl-5'-TBDMS-3'-*xylo*-2'-deoxyadenosine (**39**)



N-6-Allyloxycarbonyl-5'- TBDMS-2'-deoxyadenosine (**38**, 1.37 g, 3.05 mmol) dissolved in anhydrous DCM (10 mL) was added to a cooled (0 °C) solution of Dess Martin periodane (15% wt in DCM, 19.4 mL, 6.86 mmol) and allowed to stir for 30 mins. The reaction was allowed to warm to ambient temperature before being allowed to stir for 12 hours. Isopropanol (8 mL) was added and the reaction was

cooled to -60°C. Sodium borahydride (230 mg, 6.10 mmol) was added. After stirring for 2 hours the reaction was allowed to warm to ambient temperature and acetone added (10 mL). The resulting mixture was then poured into ethyl acetate (100 mL) and washed successively with saturated sodium hydrogen carbonate (100 mL), water (100 mL) and brine (100 mL). The organic layer was dried over $MgSO_4$ and all reaction solvents were removed in vacuo. The crude residue was purified by silica chromatography using a 2-4% methanol/DCM eluent system and the desired product was afforded as a pale yellow foam.

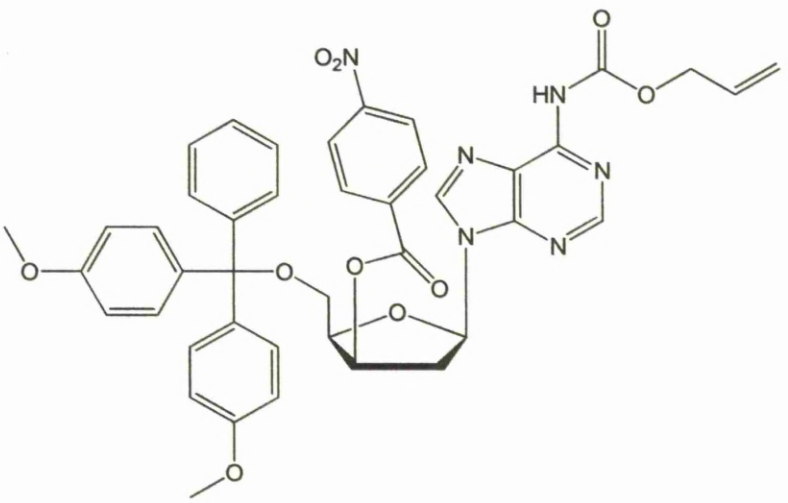
Yield: 236 mg, 0.52 mmol, 17%

^{1}H NMR (400MHz, $CDCl_3$) δ –0.02 (s, 3H, CH_3), 0.00 (s, 3H, CH_3), 0.81 (s, 9H, *t*-Bu), 2.50 (m, 1H, H2'), 2.81 (m, 1H, H2'), 3.95 (m, 2H, H5'), 4.02 (apparent q, 1H, *J*= 5.12 Hz, H4'), 4.48 (m, 1H, H3'), 4.71 (m, 2H, $C\underline{H}_2CH=CH_2$), 5.23 (apparent dq, 1H, *J*= 0.80, 10.36 Hz, $CH=C\underline{H}_2$), 5.35 (apparent dq, 1H, *J*= 1.12, 17.16 Hz, $CH=C\underline{H}_2$), 5.95 (m, 1H, $C\underline{H}=CH_2$), 6.21 (dd, 1H, *J*= 2.42, 9.06 Hz, H1'), 8.21 (s, 1H, H8), 8.54 (s, 1H, NH), 8.69 (s, 1H, H2).

^{13}C NMR (100MHz, $CDCl_3$) δ -4.9 (2 x CH_3), 18.7 ($\underline{C}(CH_3)_3$), 26.2 ($C(\underline{C}H_3)_3$), 41.3 (C2'), 62.5 (C5'), 67.1 ($\underline{C}H_2CH=CH_2$), 71.5 (C3'), 84.6 (C1'), 85.1 (C4'), 119.5 ($CH=\underline{C}H_2$), 123.2 (C5), 132.0 ($\underline{C}H=CH_2$), 143.2 (C8), 150.1 (C=O Alloc), 150.3 (C6), 151.0 (C4), 152.6 (C2),

HRMS found m/z (ES^+) 472.1974 ($[M + Na]^+$ 100%); $[C_{20}H_{31}N_5O_5 + Na]^+$ requires 472.1992.

Preparation of *N*-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*(*para*-nitrobenzoyl)-2′-deoxyadenosine (**41**)



N-6-Allyloxycarbonyl-5′- dimethoxytrityl-2′-deoxyadenosine (**40**, 2.85 g, 4.47 mmol), *p*-nitrobenzoic acid (1.49 g, 8.94 mmol) and triphenylphosphine (2.34 g, 8.94 mmol) were suspended in anhydrous THF (60 mL) and DEAD (40% in toluene, 3.89 mL, 8.94 mmol) was added dropwise. The reaction was allowed to stir at ambient temperature for 2 hours and after this time all solvents were removed in vacuo. The crude residue was purified by silica chromatography using a 0-2% methanol/chloroform eluent system with 1% triethylamine to basify the silica. The desired product was afforded as a pale yellow solid.

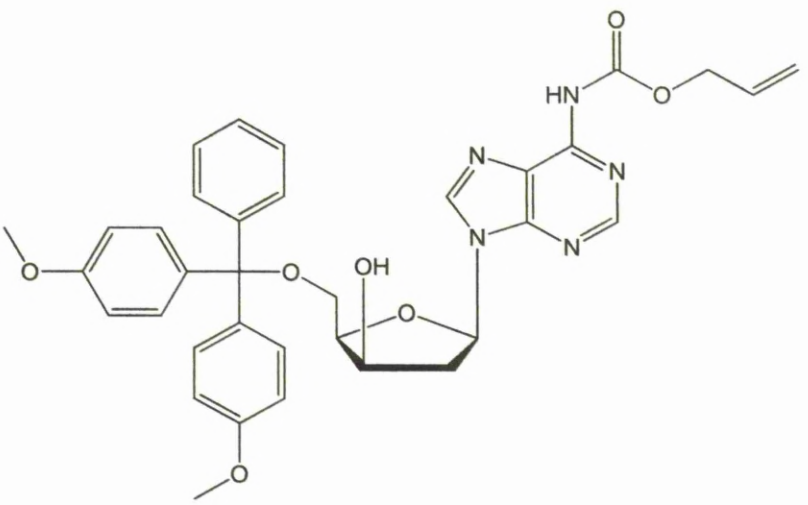
Yield: 2.90 g, 3.69 mmol, 82%

^{1}H NMR (400MHz, DMSO-d_6) δ 3.07 (m, 1H, H2′), 3.27 (m, 1H, H2′), 3.36 (m, 2H, H5′), 3.65 (s, 3H, O-CH_3), 3.67 (s, 3H, O- CH_3), 4.65 (m, 3H, $CH_2CH{=}CH_2$ + H4′), 5.23 (apparent dq, 1H, *J*= 1.69, 10.38 Hz, $CH{=}CH_2$), 5.40 (apparent dq, 1H, *J*= 1.47, 17.22 Hz, $CH{=}CH_2$), 5.89 (m, 1H, H3′), 5.97 (m, 1H, $CH{=}CH_2$), 6.53 (dd, 1H, *J*= 2.60, 7.48Hz, H1′), 6.71 (m, 4H, ArH DMT), 7.09-7.31 (m, 9H, ArH DMT), 7.69 (m, 2H, ArH PNBz), 8.23 (m, 2H, ArH PNBz), 8.44 (s, 1H, H8), 8.50 (s, 1H, H2), 10.63 (s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 37.0 (C2′), 55.2 (O-CH_3), 55.2 (O-CH_3), 61.3 (C5′), 65.5 ($CH_2CH{=}CH_2$), 74.0 (C3′), 81.4 (C4′), 84.0 (C1′), 86.0 (DMT *q*-C), 113.4 (ArC), 113.5 (ArC), 117.9 ($CH{=}CH_2$), 123.9 (ArC), 124.0 (C5), [129.1, 129.9, 130.6, 131.0, 131.7, (ArC)], 131.8 ($CH{=}CH_2$), [134.8, 135.2, 135.4 (ArC)], 142.2 (C8), 152.2 (C6), 156.9 (C6), 158.3 (C4), 158.4 (C2), 165.3 (C=O PNBz).

HRMS found m/z (ES^+) 809.2568 ([M + Na]$^+$ 100%); [$C_{42}H_{38}N_6O_{10}$ + Na]$^+$ requires 809.2547.

Preparation of *N*-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*-2′-deoxyadenosine (**42**)



N-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*(*para*-nitrobenzoyl)-2′-deoxyadenosine (**41**, 2.80 g, 3.56 mmol) and sodium methoxide (384 mg, 7.12 mmol) were dissolved in anhydrous methanol (200 mL) and allowed to stir at 0°C for 25 mins. The reaction solvents were removed in vacuo and the resulting residue redissolved in DCM (100 mL). The organic solution washed twice with water (2 x100 mL). The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude residue was purified by silica chromatography using a 0-2% methanol/DCM eluent system containing 1% triethylamine to basify the silica. The desired product was afforded as a pale yellow foam

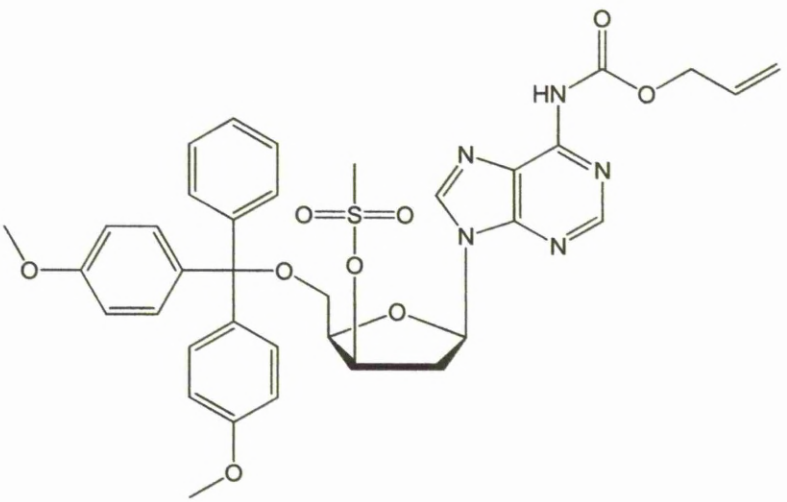
Yield: 1.85 g, 2.9 mmol, 81%

^{1}H NMR (400MHz, DMSO-d_6) δ 2.38 (m, 1H, H2′), 2.77 (m, 1H, H2′), 3.20 (m, 1H, H5′), 3.35 (m, 1H, H5′), 3.70 (s, 3H, O-CH_3), 3.71 (s, 3H, O-CH_3), 4.24 (apparent dt, 1H, *J*= 3.58, 7.42 Hz, H4′), 4.36 (apparent q, 1H, *J*= 4.28 Hz, H3′), 4.66 (m, 2H, $C\underline{H}_2CH=CH_2$), 5.24 (apparent dq, 1H, *J*= 1.50, 1052 Hz, $CH=C\underline{H}_2$), 5.42 (apparent dq, 1H, *J*= 1.72, 17.22 Hz, $CH=C\underline{H}_2$), 4.36 (d, 1H, *J*= 4.36 Hz, 3′-OH), 5.98 (m, 1H, $C\underline{H}=CH_2$), 6.48 (dd, 1H, *J*= 1.72, 8.48 Hz, H1′), 6.80 (m, 4H, ArH), 7.24 (m, 7H, ArH), 7.39 (m, 2H, ArH), 8.46 (s, 1H, H8), 8.65 (s, 1H, H2), 10.67 (s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 40.9 (C2′), 55.3 (O-CH_3), 63.5 (C5′), 65.5 ($\underline{C}H_2CH=CH_2$), 69.6 (C3′), 83.5 (C1′), 84.3 (C4′), 85.8 (DMT *q*-C), 113.3 (ArC), 113.4 (ArC), 117.9 ($CH=\underline{C}H_2$), 126.9 (C5), 131.8 ($\underline{C}H=CH_2$), [133.2, 135.8, 136.0 (ArC)], 142.9 (C8), [145.3, 149.9, 151.6 (ArC)], 151.8 (C=O Alloc), 152.2 (C6), 156.9 (C4), 158.3 (C2).

HRMS found m/z (ES$^+$) 660.2431 ([M + Na]$^+$ 100%); [$C_{35}H_{35}N_5O_7$ + Na]$^+$ requires 660.2434.

Preparation of *N*-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*(mesyl)-2′-deoxyadenosine (**43**)



N-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*-2′-deoxyadenosine (**42**,1.85 g, 2.90 mmol) and dimethylaminopyridine (1.77g, 14.55 mmol) were dissolved in anhydrous pyridine (70 mL) and methanesulphonyl chloride (676 µl, 8.73 mmol) was added dropwise. The reaction was allowed to stir at ambient temperature for 2 hours. After this time all reaction solvents were removed in vacuo. The crude residue was purified using a 0-3% methanol/chloroform eluent system and the desired product was afforded as a pale yellow foam.

Yield: 1.30 g, 1.90 mmol, 65%

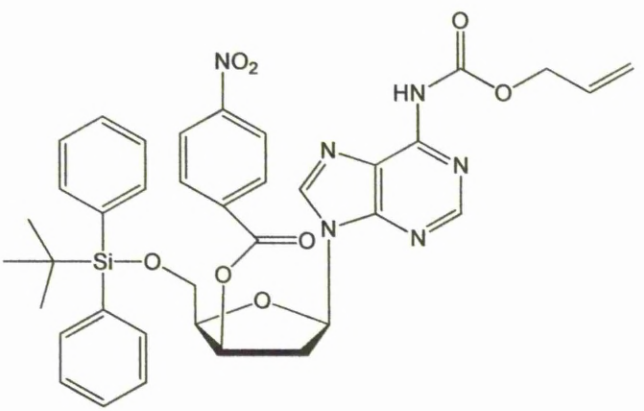
^{1}H NMR (400MHz, DMSO-d_6) δ 2.87 (m, 1H, H2′), 3.07 (m, 4H, $SO_2C\underline{H}_3$ + H2′), 3.21 (dd, 1H, *J*= 4.06, 10.03 Hz, H5′), 3.39 (dd, 1H, *J*= 7.36, 10.03 Hz, H5′), 3.72 (s, 6H, 2 x O-CH_3), 4.50 (m, 1H, H4′), 4.66 (m, 2H, $C\underline{H}_2CH=CH_2$), 5.24 (apparent dq, 1H, *J*= 1.47, 10.54 Hz, $CH=C\underline{H}_2$), 5.42 (apparent dq, 1H, *J*= 1.65, 17.24 Hz, $CH=C\underline{H}_2$), 5.48

(m, 1H, H3′), 5.98 (m, 1H, C<u>H</u>=CH_2), 6.55 (dd, 1H, *J*= 2.40, 7.92 Hz, H1′), 6.83 (m, 4H, ArH), 7.26 (m, 7H, ArH), 7.40 (m, 2H, ArH), 8.26 (s, 1H, H8), 8.32 (s, 1H, H2), 10.71 (br s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 37.8 (C2′), 38.6 (SO_2<u>C</u>H_3), 55.3 (O-<u>C</u>H_3), 60.8 (C5′), 65.6 (<u>C</u>H_2CH=CH_2), 80.2 (C3′), 81.8 (C4′), 83.0 (C1′), 113.1 (ArC), 113.4 (ArC), 117.9 (CH=<u>C</u>H_2), 124.3 (C5), [129.2, 130.0, 130.1 (ArC)], 131.8 (<u>C</u>H=CH_2), [133.2, 133.5, 135.5 (ArC)], 144.9 (C=O Alloc), 141.6 (C8), 149.7 (C6), 150.0 (C4), 156.9 (C2).

HRMS found m/z (ES$^+$) 738.2236 ([M + Na]$^+$ 100%); [$C_{36}H_{37}N_5O_9$ + Na]$^+$ requires 738.2210.

<u>Preparation of *N*-6-Allyloxycarbonyl-5′-TBDPS-3′-*xylo*(*para*-nitrobenzoyl)-2′-deoxyadenosine (**46**)</u>



N-6-Allyloxycarbonyl-5′-TBDPS-2′-deoxyadenosine (**45**, 6.14 g, 10.7 mmol), *p*-nitrobenzoic acid (3.57 g, 21.4 mmol) and triphenylphosphine (5.61 g, 21.4 mmol) were dissolved in anhydrous tetrahydrofuran (100 mL) under a nitrogen atmosphere. A solution of DEAD (40% in toluene, 9.31 mL, 21.4 mmol) was added dropwise with stirring and the reaction was allowed to stir at ambient temperature for 3 hours. All solvents were removed from the reaction in vacuo and the crude residue was purified by silica chromatography using 100% ethyl acetate. The product was afforded as an off white foam.

Yield: 5.60 g, 7.75 mmol, 72%

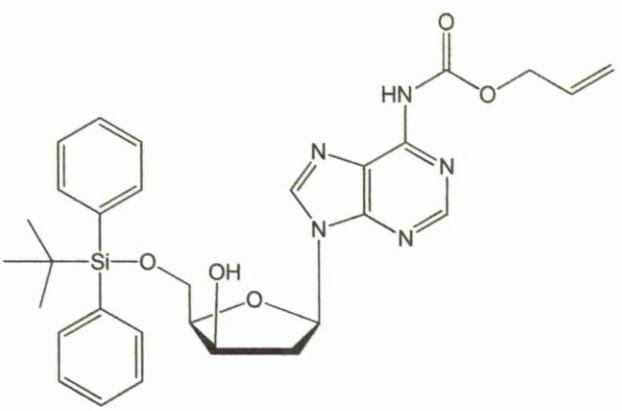
^{1}H NMR (400MHz, $CDCl_3$) δ 0.92 (s, 9H, *t*-Bu), 3.00 (m, 1H, H2′), 3.10 (m, 1H, H2′), 4.08 (m, 2H, H5′), 4.55 (m, 1H, H4′), 4.73 (m, 2H, <u>C</u>H_2CH=CH_2), 5.26 (apparent dq,

1H, *J*= 1.17, 10.44 Hz, CH=C$\underline{H}_2$), 5.38 (apparent dq, 1H, *J*= 1.46, 17.18 Hz, CH=C$\underline{H}_2$), 5.95 (m, 2H, C$\underline{H}$=CH_2 + H3′), 6.50 (dd, 1H, *J*= 1.86, 7.18 Hz, H1′), 7.19-7.75 (m, 12H, ArH),8.14 (m, 2H, ArH), 8.21 (s, 1H, H8), 8.61 (s, 1H, H2), 8.90 (br s, 1H , NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 19.4 (C($\underline{C}H_3)_3$), 27.0 (C($\underline{C}H_3)_3$), 39.4 (C2′), 60.8 (C5′), 67.1 ($\underline{C}H_2CH$=CH_2), 73.7 (C3′), 84.0 (C1′), 85.3 (C4′), 119.5 (CH=$\underline{C}H_2$), 122.8 (C5), [128.2, 129.0, 130.3, 131.4 (ArC)], 132.0 ($\underline{C}H$=CH_2), 132.9 (ArC), 134.6 (ArC), 141.0 (C8), 149.0 (C=O Alloc), 153.0 (C6), 157.2 (C4), 163.6 (C2), 171.6 (C=O PNBz),

HRMS found m/z (ES$^+$) 745.2424 ([M + Na]$^+$ 100%); [$C_{37}H_{38}N_6O_8Si$ + Na]$^+$ requires 745.2418.

Preparation of *N*-6-Allyloxycarbonyl-5′-TBDPS-3′-*xylo*-2′-deoxyadenosine (**47**)



N-6-Allyloxycarbonyl-5′-TBDPS-3′-*xylo*(*para*-nitrobenzoyl)-2′-deoxyadenosine (**46**, 5.60 g, 7.75 mmol) was dissolved in anhydrous methanol (350 mL) and cooled to 0°C in an ice bath. Sodium methoxide powder (419 mg, 7.76 mmol) was then added and the solution allowed to stir at 0°C for 10 minutes. The reaction solvents were removed in vacuo and the residue redissolved in DCM (150 mL) and washed with water (3 x 100 mL). The organic layer was dried over $MgSO_4$ and all solvent was removed in vacuo. The crude residue was purified by silica chromatography using a 1-4% methanol/DCM eluent system. The product was afforded as a white foam.

Yield: 3.29 g, 5.74 mmol, 74%

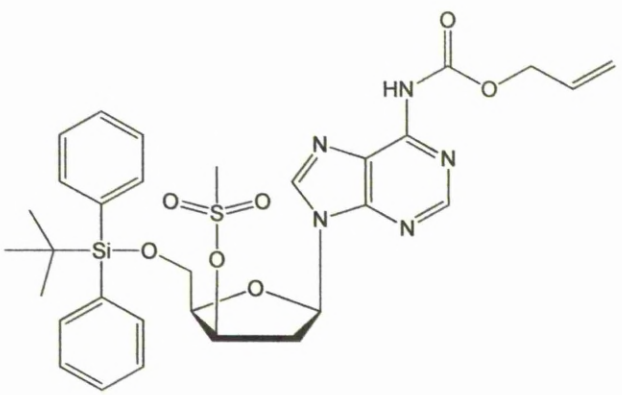
^{1}H NMR (400MHz, $CDCl_3$) δ 1.04 (s, 9H, *t*-Bu), 2.54 (m, 1H, H2′), 2.86 (m, 1H, H2′), 3.99-4.14 (m, 3H, H5′ + H4′), 4.52 (apparent q, 1H, *J*= 2.89 Hz, H3′), 4.77 (apparent dt, 2H, *J*= 1.28, 5.84 Hz, $\underline{C}H_2CH$=CH_2), 5.29 (apparent dq, 1H, *J*= 1.16, 10.44 Hz,

CH=C$\underline{H}_2$), 5.41 (apparent dq, 1H, *J*= 1.44, 17.16 Hz, CH=C$\underline{H}_2$), 5.99 (m, 1H, C$\underline{H}$=CH$_2$), 6.25 (dd, 1H, *J*= 2.28, 9.04 Hz, H1'), 7.36 (m, 6H, ArH), 7.63 (m, 4H, ArH), 8.28 (s, 1H, H8), 8.70 (s, 1H, H2).

^{13}C NMR (100MHz, $CDCl_3$) δ 19.5 ($\underline{C}(CH_3)_3$), 27.2 (C($\underline{C}H_3)_3$), 41.2 (C2'), 63.1 (C5'), 67.1 ($\underline{C}H_2CH=CH_2$), 71.3 (C3'), 84.7 (C1'), 85.3 (C4'), 119.5 (CH=$\underline{C}H_2$), 128.0 (ArC), 123.3 (C5), [128.0, 130.1, 130.2 (ArC)], 132.0 ($\underline{C}H=CH_2$), 133.3 (ArC), 135.9 (ArC), 143.3 (C8), 150.2 (C=O Alloc), 150.2 (C6), 151.1 (C4), 152.5 (C2).

HRMS found m/z (ES$^+$) 596.2281 ([M + Na]$^+$ 100%); [$C_{30}H_{35}N_5O_5Si$ + Na]$^+$ requires 596.2305.

Preparation of *N*-6-Allyloxycarbonyl-5'-TBDPS -3'-*xylo*(mesyl)-2'-deoxyadenosine (**48**)



N-6-Allyloxycarbonyl-5'-TBDPS-3'-*xylo*-2'-deoxyadenosine (**47**, 2.26 g, 3.94 mmol) and dimethylaminopyridine (2.40 g, 19.70 mmol) were dissolved in anhydrous pyridine (150mL) under nitrogen. Methanesulphonyl chloride (915 µl, 11.82 mmol) was then added dropwise and the reaction allowed to stir at ambient temperature for 2 hours. All reaction solvents were then removed in vacuo and the crude residue purified by silica chromatography using a 0-3% methanol/DCM eluent system. This afforded the product as an off-white foam.

Yield: 2.10 g, 3.22 mmol, 82%

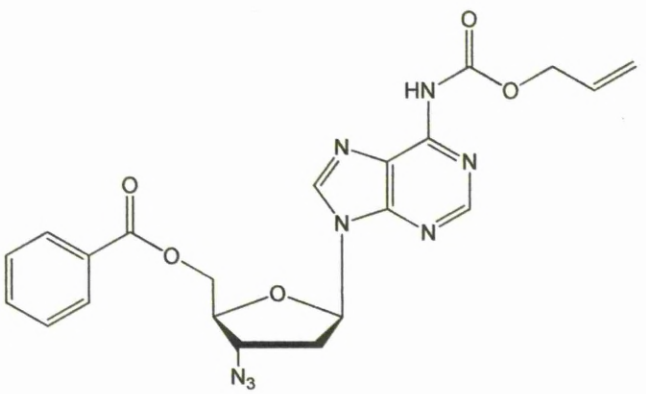
^{1}H NMR (400MHz, $CDCl_3$) δ 1.08 (s, 9H, *t*-Bu), 2.87 (s, 3H, $SO_2C\underline{H}_3$), 2.95 (m, 2H, H2'), 4.01 (m, 2H, H5'), 4.28 (m, 1H, H4'), 4.76 (apparent dt, 2H, *J*= 1.30, 5.80Hz, $C\underline{H}_2CH=CH_2$), 5.28 (apparent dq, 1H, *J*= 1.17, 9.22 Hz, CH=C$\underline{H}_2$), 5.40 (apparent dq, 1H, *J*= 1.45,17.18 Hz, CH=C$\underline{H}_2$), 5.45 (apparent q, 1H, *J*= 2.97 Hz, H3'), 5.97

(m, 1H, C<u>H</u>=CH_2), 6.52 (apparent t, 1H, *J*= 5.10 Hz, H1'), 7.41 (m, 6H, ArH), 7.67 (m, 4H, ArH), 8.18 (s, 1H, H8), 8.44 (br s, 1H, NH), 8.73 (s, 1H, H2).

^{13}C NMR (100MHz, $CDCl_3$) δ 19.5 (<u>C</u>(CH_3)$_3$), 27.2 (C(<u>C</u>H_3)$_3$), 38.9 (SO_2<u>C</u>H_3), 40.2 (C2'), 61.2 (C5'), 67.0 (<u>C</u>H_2CH=CH_2), 78.5 (C3'), 83.4 (C1'), 83.6 (C4'),119.4 (CH=<u>C</u>H_2), 122.4 (C5), 128.3 (ArC), 132.9 (<u>C</u>H=CH_2), [133.9, 135.9, 136.3, (ArC)], 141.0 (C8), 150.2 (C=O Alloc), 151.1 (C6), 151.3 (C4),153.3 (C2).

HRMS found m/z (ES$^+$) 674.2069 ([M + Na]$^+$ 100%); [$C_{31}H_{37}N_5O_7SiS$ + Na]$^+$ requires 674.2081

<u>Preparation of *N*-6-Allyloxycarbonyl-5'-benzoyl-3'-azido-2', 3'-dideoxyadenosine (**34**)</u>



N-6-Allyloxycarbonyl-5'-benzoyl-3'-*xylo*-2'-deoxyadenosine (**31**, 1.18 g, 2.69 mmol) was dissolved in anhydrous DCM (30 mL) and pyridine (3 mL) and cooled to -30°C. Triflic anhydride (680 µl) was then added dropwise. The reaction was allowed to warm to ambient temperature over 30mins. A solution of lithium azide (1.32 g, 27.0 mmol) in anhydrous DMF (10 mL) was then added and the reaction allowed to stir at ambient temperature for 2 hours. Water (50 mL) was added to the reaction and the two layers were separated. The organic layer was washed with water (50 mL) and then dried over $MgSO_4$. The crude residue was purified by silica chromatography using a 0-2% methanol/DCM eluent system the desired product was afforded as a yellow solid.

Yield: 800 mg, 1.72 mmol, 65%

^{1}H NMR (400MHz, $CDCl_3$) δ 2.67 (m, 1H, H2'), 3.25 (m, 1H, H2'), 4.31 (apparent q, 1H, *J*= 4.80 Hz, H4'), 4.57(dd, 1H, *J*= 4.56, 12.20 Hz, H5'), 4.68 (dd, 1H, *J*= 4.16, 12.20 Hz, H5'), 4.77 (m, 3H, C<u>H</u>$_2$CH=CH_2 + H3'), 5.29 (apparent dq, 1H, *J*= 1.18,

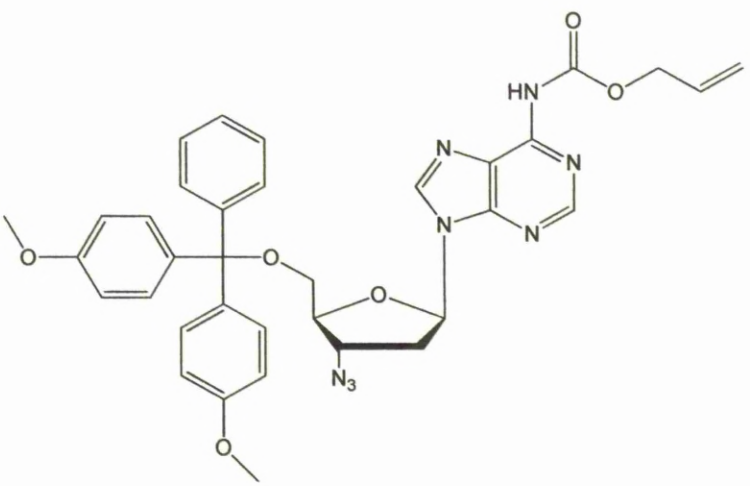
10.46 Hz, CH=C<u>H</u>$_2$), 5.40 (apparent dq, 1H, *J*= 1.44, 17.18 Hz, CH=C<u>H</u>$_2$), 5.99 (m, 1H, C<u>H</u>=CH$_2$), 6.34 (dd, 1H, *J*= 5.08, 7.00 Hz, H1′), 7.39 (m, 2H, *m*-ArH), 7.55 (m, 1H, *p*-ArH), 7.93 (m, 2H, *o*-ArH), 8.12 (s, 1H, H8), 8.69 (s, 1H, H2), 8.85 (br s, 1H, NH).

^{13}C NMR (400MHz, $CDCl_3$) δ 37.4 (C2′), 61.2 (C3′), 63.8 (C5′), 67.0 (<u>C</u>H$_2$CH=CH$_2$), 82.9 (C4′), 85.2 (C1′), 119.5 (CH=<u>C</u>H$_2$), 123.1 (C5), 128.9 (*m*-ArC), 129.5 (*ipso*-ArC), 129.9 (*o*-ArC), 132.1 (<u>C</u>H=CH$_2$), 133.8 (*p*-ArC), 142.2 (C8), 150.0 (C=O Alloc), 150.9 (C6), 151.2 (C4), 153.2 (C2), 166.4 (C=O Bz).

HRMS found m/z (ES$^+$) 487.1456 ([M + Na]$^+$ 100%); [$C_{21}H_{20}N_8O_5$ + Na]$^+$ requires 487.1454

IR *v* max/ cm^{-1} = 1074.16, 1211.08, 1270.86, 1469.49, 1616.06, 1718.26, 1756.76, 2090.46, 3313.11

<u>Preparation of *N*-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-azido-2′, 3′-deoxyadenosine (**44**)</u>

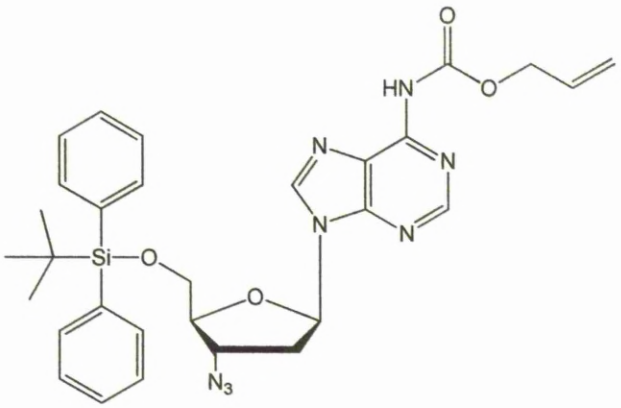


N-6-Allyloxycarbonyl-5′-dimethoxytrityl-3′-*xylo*(mesyl)-2′-deoxyadenosine (**43**, 1.30 g, 1.82 mmol) and lithium azide (552 mg, 10.9 mmol) were dissolved in anhydrous DMF (40 mL) and allowed to stir at 60°C for 7 hours. Water (1 mL) was added and after cooling to ambient temperature all solvents were removed in vacuo. The crude residue was purified by silica chromatography using a 0-2% methanol/ chloroform eluent system. The product was afforded as a yellow oil

Yield: 720 mg, 1.08 mmol, 72%

HRMS found m/z (ES^+) 685.2495 ($[M + Na]^+$ 100%); $[C_{35}H_{34}N_8O_6 + Na]^+$ requires 685.2499.

Preparation of *N*-6-Allyloxycarbonyl-5′-TBDPS -3′-azido-2′, 3′-deoxyadenosine (**49**)



N-6-Allyloxycarbonyl-5′-TBDPS -3′-*xylo*(mesyl)-2′-deoxyadenosine (**48**, 1.48 g, 2.27 mmol) and lithium azide (555 mg, 11.4 mmol) were suspended in anhydrous DMF (40 mL) under a nitrogen atmosphere and heated to 70°C. The reaction was allowed to stir for 2 hours and then removed from the heat and allowed to cool. The reaction solvents were removed in vacuo and the crude residue redissolved in DCM (100 mL). The organic solution was washed with successively a saturated sodium hydrogen carbonate solution (100 mL), water (100 mL) and brine (100 mL). The organic layer was dried over $MgSO_4$ and evaporated to dryness. The crude residue was purified by silica chromatography using a 2% methanol/DCM eluent system. The product was afforded as a white foam.

Yield: 640 mg, 1.07 mmol, 47%

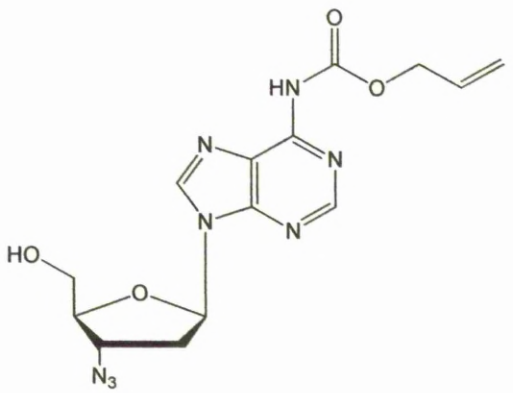
^{1}H NMR (400MHz, $CDCl_3$) δ 0.98 (s, 9H, *t*-Bu), 2.51 (m, 1H, H2′), 2.91 (m, 1H, H2′), 3.74 (dd, 1H, *J*= 3.76, 11.44 Hz, H5′), 3.89 (dd, 1H, *J*= 4.64, 11.44 Hz, H5′), 4.00 (apparent q, 1H, *J* = 4.24 Hz, H4′), 4.49 (m, 1H, H3′), 4.69 (apparent dt, 2H, *J*= 1.31, 5.83Hz, $C\underline{H}_2CH=CH_2$), 5.21 (apparent dq, 1H, *J*= 1.23, 11.54 Hz, $CH=C\underline{H}_2$), 5.33 (apparent dq, 1H, *J*= 1.44, 16.44Hz, $CH=C\underline{H}_2$), 5.91 (m, 1H, $C\underline{H}=CH_2$), 6.29 (apparent t, 1H, *J*= 6.20 Hz, H1′), 7.26-7.38 (m, 6H, ArH), 7.55 (m, 4H, ArH), 8.09 (s, 1H, H8), 8.32 (br s, 1H, NH), 8.62 (s, 1H, H2).

^{13}C NMR (100MHz, $CDCl_3$) δ 19.6 ($\underline{C}(CH_3)_3$), 27.2 ($C(\underline{C}H_3)_3$), 37.7 (C2′), 60.9 (C3′), 63.6 (C5′), 67.0 ($\underline{C}H_2CH=CH_2$), 84.9 (C1′), 85.3 (C4′), 119.4 ($CH=\underline{C}H_2$), 122.9 (C5),

128.1 (ArC), 128.2 (ArC), 132.1 (CH=CH_2), [132.8, 133.0, 135.8 (ArC)], 141.6 (C8), 149.8 (C=O Alloc), 151.1 (C6), 151.0 (C4), 153.2 (C2).

HRMS found m/z (ES^+) 621.2385 ($[M + Na]^+$ 100%); $[C_{30}H_{34}N_8O_4Si + Na]^+$ requires 621.2370.

Preparation of *N*-6-allyloxycarbonyl-3′-azido-2′, 3′-dideoxyadenosine (**35**)



Removal of Benzoyl protecting group

N-6-Allyloxycarbonyl-5′-benzoyl-3′-azido-2′-deoxyadenosine (**34**, 800 mg, 1.72 mmol) was dissolved in a dioxane: water: methanol mix (5:4:1, 30 mL) and cooled in ice. 1M (aq) Sodium hydroxide (3.5 mL) was then added and the reaction allowed to stir for 30 mins. Amberlite™ ion exchange resin (1 g) was added to the reaction mixture to achieve neutral pH. The resin was then removed by filteration and all solvents removed in vacuo. The crude residue was redissolved in DCM (20 mL), washed with water (20 mL) and the organic layer dried over $MgSO_4$. The reaction solvents were removed in vacuo to afford a yellow foam.

Yield: 360 mg, 1.00 mmol, 58%

Removal of TBDPS protecting group

N-6-Allyloxycarbonyl-5′-TBDPS -3′-azido-2′, 3′-deoxyadenosine (**49**, 600 mg, 1.00 mmol) was dissolved in anhydrous THF (7 mL) and TBAF (1.20 mL, 1.20 mmol) was added dropwise. The reaction was allowed to stir at ambient temperature for 2 hours then all solvents were removed in vacuo. The residue was redissolved in dichloromethane (50 mL) and washed with water (2 x 50 mL). The organic layer was dried over $MgSO_4$ and all solvents were removed in vacuo. The crude residue was

purified by silica chromatography using a 2% methanol/DCM eluent system. The product was afforded as a white foam.

Yield: 201 mg, 0.56 mmol, 55%

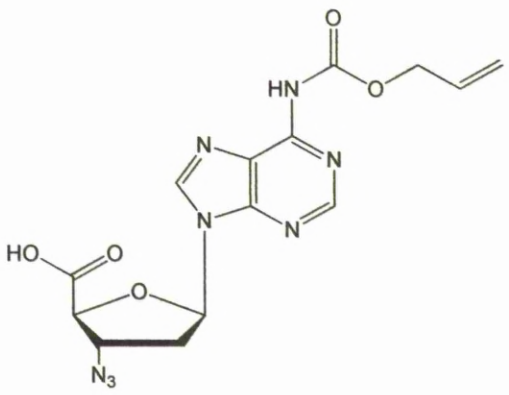
^{1}H NMR (400MHz, $CDCl_3$) δ 2.47 (m, 1H, H2′), 3.11 (m, 1H, H2′), 3.82 (m, 1H, H5′), 4.03 (m, 1H, H5′), 4.23 (m, 1H, H4′), 4.63 (apparent dt, 1H, *J*= 2.29, 6.09 Hz, H3′), 4.76 (m, 2H, $C\underline{H}_2CH=CH_2$), 5.29 (m, 1H, $CH=C\underline{H}_2$), 5.40 (m, 1H, $CH=C\underline{H}_2$), 5.99 (m, 1H, $C\underline{H}=CH_2$), 6.32 (dd, 1H, *J*= 5.82, 8.34 Hz, H1′), 8.21 (s, 1H, H8), 8.72 (s, 1H, H2), 9.63 (br s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 38.2 (C2′), 62.5 (C3′), 63.5 (C5′), 67.0 ($\underline{C}H_2CH=CH_2$), 86.8 (C4′), 87.2 (C1′), 119.6 ($CH=\underline{C}H_2$), 123.6 (C5), 132.0 ($\underline{C}H=CH_2$), 142.8 (C8), 150.5 (C=O Alloc), 150.6 (C6), 151.4 (C4), 152.6 (C2).

HRMS found m/z (ES^+) 383.1188 ($[M + Na]^+$ 100%); $[C_{14}H_{16}N_8O_4 + Na]^+$ requires 383.1192

IR *v* max/ cm^{-1} = 1099.23, 1211.08, 1465.63, 1585.20, 1616.06,1754.90, 2103.96, 2942.84, 3251.40.

Preparation of *N*-6-Allyloxycarbonyl-3′-azido-2′, 3′-dideoxyadenosine-5′-carboxylic acid (**36**)



N-6-Allyloxycarbonyl-3′-azido-2′-deoxyadenosine (**35**, 360 mg, 1.00 mmol), TEMPO (31.0 mg, 0.20 mmol) and BAIB (730 mg, 2.20 mmol) were suspended in an acetonitrile: water mixture (1:1, 4mL) and allowed to stir at ambient temperature for 4 hours. All solvents were removed in vacuo and the product was redissolved in a minimum amount of methanol. The pure product was obtained by precipitation from

the methanol solution using diethyl ether. The desired product was afforded as an off white powder.

Yield: 270 mg, 0.72 mmol, 72%

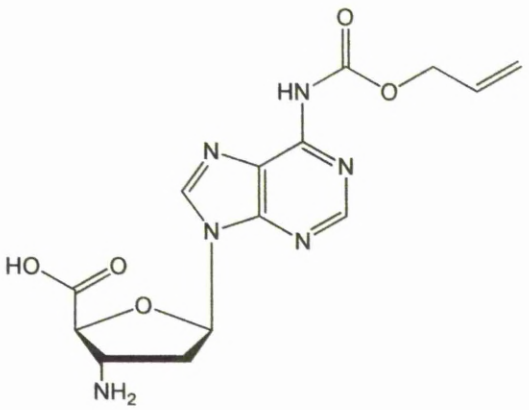
^{1}H NMR (400MHz, MeOD) δ 1.48 (m, 1H, H2′), 1.69 (m, 1H, H2′), 3.45 (d, 1H, *J*= 3.40 Hz, H4′), 3.57 (apparent dt, 2H, J= 1.41, 5.89 Hz, C$\underline{H}_2$CH=CH$_2$), 3.70 (m, 1H, H3′), 4.09 (apparent dq,1H, *J*= 1.32, 10.50Hz, CH=C$\underline{H}_2$), 4.25 (apparent dq, 1H, *J*= 1.57, 17.24 Hz, CH=C$\underline{H}_2$), 4.86 (m, 1H, C$\underline{H}$=CH$_2$), 5.42 (apparent t, 1H, *J*=6.32 Hz, H1′), 7.41 (s, 1H, H8), 7.55 (s, 1H, H2).

^{13}C NMR (400MHz, MeOD) δ 36.9 (C2′), 64.1 (C3′), 65.5 (C5), 65.9 ($\underline{C}$H$_2$CH=CH$_2$), 82.6 (C4′), 85.3 (C1′), 117.2 (CH=$\underline{C}$H$_2$), 132.2 ($\underline{C}$H=CH$_2$), 142.7 (C2), 149.5 (C6), 151.7 (C4), 152.0 (C8).

HRMS found m/z (ES$^+$) 375.1165 ([M + H]$^+$ 100%); [$C_{14}H_{15}N_8O_5$]$^+$ requires 375.1165.

IR *v* max/cm^{-1}= 1029.80, 1095.37, 1222.65, 1326.79, 1415.49, 1554.34, 1581.34, 1727.91, 1766.48, 2102.03, 2989.12, 3185.83.

Preparation of *N*-6-Allyloxycarbonyl-3′-amino-2′, 3′-dideoxyadenosine-5′-carboxylic acid (**37**)

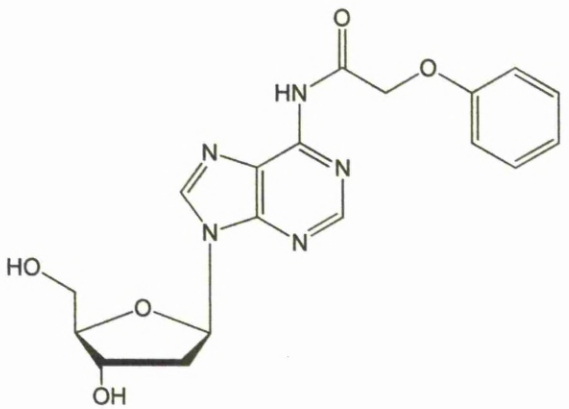


N-6-Allyloxycarbonyl-3′-azido-2′-deoxyadenosine-5′-carboxylic acid (**36**, 210 mg, 0.56 mmol) was dissolved in an anhydrous 15% triethylamine/pyridine solution (15 mL) and cooled to 0°C. Hydrogen sulphide was bubbled through the mixture for 10 mins then the reaction vessel sealed and allowed to stir in an ice bath for a further

30 mins. All reaction solvents were removed in vacuo to afford a brown solid. Kept as crude.

HRMS found m/z (ES^+) 371.1070 ($[M + H]^+$ 100%); $[C_{14}H_{16}N_6O_5]^+$ requires 371.1080.

Preparation of *N*-6-Phenoxyacetyl-2′-deoxyadenosine (**26**)



2′-Deoxyadenosine (0.50 g, 1.99 mmol) was co-evaporated with anhydrous pyridine (3 x 10mL) and dried thoroughly under vacuum. The white solid was then suspended in anhydrous pyridine (8 mL) and trimethylsilylchloride (1.77 mL, 14.0 mmol) was added dropwise. The reaction was stirred at ambient under nitrogen for 2 hours. In a separate flask triazole (192 mg, 2.79 mmol) was dissolved in anhydrous pyridine (5 mL) and cooled to 5°C in an ice bath. Phenoxyacetal chloride (385 µL, 2.79 mmol) was then added dropwise and the reaction allowed to stir at ambient temperature under nitrogen for 10 minutes. This was then added to the solution containing the protected nucleoside and the reaction allowed to stir at ambient temperature under nitrogen for 18 hours. Water (1 mL) was added to the reaction mixture after cooling in an ice bath. The reaction was stirred for a further hour. All reaction solvents were then removed in vacuo, the residue was then redissolved in water (10 mL) and extracted with DCM (3 x 10 mL). The combined organics were dried over $MgSO_4$ and the solvents removed in vacuo. The crude residue was purified by silica chromatography using a 5-10% methanol/DCM eluent system. The product was afforded as a white solid.

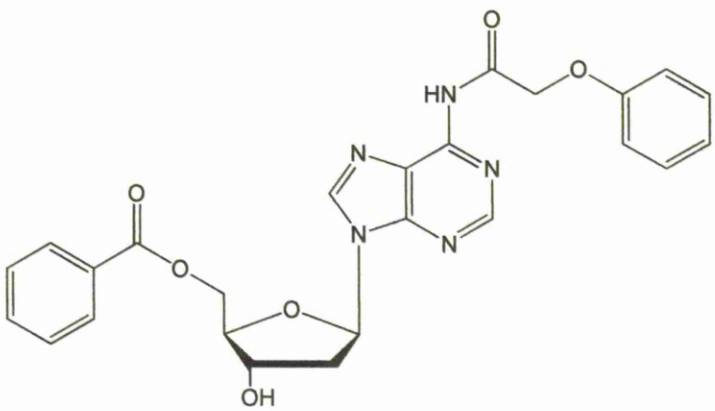
Yield: 220 mg, 0.57 mmol, 30%

^{1}H NMR (400MHz, DMSO-d_6) δ 2.35 (m, 1H, H2′), 2.77 (m, 1H, H2′), 3.58 (m, 2H, H5′), 3.90 (m, 1H, H4′), 4.44 (m, 1H, H3′), 5.03 (s, 2H, C(O)C$\underline{H_2}$OPh), 5.37 (br s, 1H, 3′-OH), 6.46 (apparent t, 1H, *J*= 6.72 Hz, H1′), 6.96 (m, 3H, ArH PAC), 7.31 (m, 2H, ArH PAC), 8.68 (s, 1H, H8), 8.70 (s, 1H, H2) 10.93 (br s, 1H, NH).

^{13}C NMR (100MHz, DMSO-d_6) δ 39.6 (C2′), 61.9 (C5′), 67.5 (C(O)$\underline{C}H_2$OPh), 71.0 (C3′), 84.1 (C1′), 88.3 (C4′), 114.9 (ArC), 121.4 (ArC), 123.7 (C5), 129.8 (ArC), 143.3 (C8), 151.8 (C6), 151.9 (C4), 158.1 (C2), 167.7 (C=O).

HRMS found m/z (ES$^+$) 408.1289 ([M + Na]$^+$ 100%); [$C_{18}H_{19}N_5O_5$ + Na]$^+$ requires 480.1284.

Preparation of *N*-6-Phenoxyacetyl-5′-benzoyl-2′-deoxyadenosine (**28**)



N-6-Phenoxyacetyl-2′-deoxyadenosine (220 mg, 0.56 mmol) was dissolved in anhydrous pyridine (1.5 mL) under a nitrogen atmosphere and cooled in an ice bath. Benzoyl chloride (0.07 mL, 0.56 mmol) in anhydrous pyridine (0.5 mL) was added to the stirred solution over a period of 30 mins and the reaction was then stirred in ice for a further 10 minutes. Water (0.1 mL) and ammonia (1 drop) were then added to the reaction mixture and stirring continued for 10 minutes. All solvents were removed in vacuo and the crude residue was redissolved in DCM (10 mL). The organic solution was then washed with water (10 mL) and dried over $MgSO_4$. The solution was then evaporated to dryness in vacuo. The crude residue was purified by silica chromatography using a 0-4% Methanol/DCM eluent system. The product was afforded as a white powdery solid.

Yield: 75 mg, 0.15 mmol, 26%

^{1}H NMR (400MHz, $CDCl_3$) δ 2.66 (m, 1H, H2′), 3.03 (m, 1H, H2′), 4.35 (m, 1H, H4′), 4.64 (m, 2H, H5′), 4.86 (m, 3H, C(O)C$\underline{H}_2$OPh + H3′), 6.46 (apparent t, 1H, *J*= 6.28 Hz, H1′), 7.05 (m, 3H, ArH), 7.32-7.62 (m, 5H, ArH), 7.96 (m, 2H, ArH), 8.16 (s, 1H, H8), 8.69 (s, 1H, H2), 9.41 (br s, 1H, NH).

^{13}C NMR (100MHz, $CDCl_3$) δ 40.1 (C2′), 64.1 (C5′), 68.4 (C(O)$\underline{C}H_2$OPh), 72.0 (C3′), 85.2 (C1′ + C4′), 115.3 (ArC), 122.8 (C5), 128.9 (ArC), 130.0 (ArC), 130.2 (ArC), 133.8 (ArC), 142.4 (C8), 150.9 (C6), 151.8 (C4), 152.9 (C2).

HRMS found m/z (ES^+) 490.1710 ($[M + Na]^+$ 100%); $[C_{23}H_{25}N_5O_6 + Na]^+$ requires 490.1703.

4.3 Solid-phase peptide synthesis

Techniques and reagents

High Performance Liquid Chromatography (HPLC)

HPLC was performed on an automated Glison HPLC system equipped with an autoinjector, a photodiode array detector and a dual hydraulic pump. Chromatographic data was handled using UniPoint version 3.0 and absorptions measured at 254nm.

Separations were performed on a reverse-phase 250mm x 4.6mm Gemini 5u C18 110Å column supplied by Phenomenex®. A gradient of 0-60% MeCN/[40% MeCN/60% 0.1M Trithethylammonium bicarbonate (TEAB)] was used for the thymidine tripeptide (**50**) over a period of 45 minutes.

UV spectroscopy

Spectra were recorded at the stated wavelengths on a Hewlett Packard 8452A spectrodiode array in 1cm quartz cuvettes.

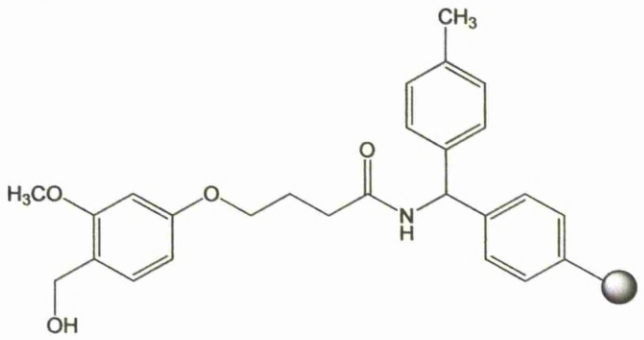
Reagents

4-Hydroxymethyl-3-methoxyphenoxybutyric acid-4-methylbenzhydrylamine (HMPB-MBHA) resin (0.82 mmol/g loading) was purchased from Novabiochem. *O*-(Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) was provided by Fluka through Sigma Aldrich.

TNBS Test:

The TNBS test was performed using the following reagents:
Reagent A: 10% v/v DIPEA/DMF
Reagent B: 1% v/v TNBS/DMF
Approximately 10 resin beads were extracted from the drained resin bed and placed in a microcentrifuge tube. One drop of each reagent solution was then added to the beads and the tube was left to stand for 5 minutes. The colour of the beads was noted as; red - positive, free amino groups present, successful deprotection. Yellow - negative, no free amino groups, unsuccessful deprotection or successful coupling.

General procedure for solid-phase synthesis of 5′-carboxylic acid peptides:
Couplings were performed on HMPB-MBHA resin (210 mg, 0.17 mmol loading) and initial coupling was carried out at 25% capacity.



Coupling	Component	Quantity
1	Fmoc amino acid	20.0 mg, 0.04 mmol
	0.2 M HBTU in DMAc	240 µL, 0.05 mmol
	DIPEA	8.30 µL, 0.05 mmol
2+	Fmoc amino acid	50.0 mg, 0.11 mmol
	0.2 M HBTU in DMAc	525 µL, 0.11 mmol
	DIPEA	18.0 µL, 0.11 mmol

Table 1: coupling conditions

Swelling of the resin and coupling of first nucleoside monomer residue:
The dry HMPB-MBHA resin was swollen overnight in three times the resin bed volume of DCM. The swelling solvent was removed and the resin washed with DMAc (x3). The thymidine derived β-amino acid (**3**, 20.0 mg, 0.04 mmol) was dissolved in 0.2 M HBTU/DMAc (240 µL, 0.05 mmol) and DIPEA (8.30 µL, 0.05 mmol) and the solution was allowed to pre-activate for 60 seconds. This solution was then added to the resin along with a further aliquot of DMAc (150 µL) and agitated on an orbital shaker for 5 hours. The coupling solution was then removed and the resin washed with DMAc (5 x 0.25 mL).

Capping unreacted sites:
Excess capping solution (Ac_2O, DMAc, collidine, 1:8:1) was added to cover the resin and the reaction vessel was agitated on an orbital shaker for 10 minutes. The resin was then washed thoroughly with DMAc (5 x 0.25 mL).

Removal of then-terminal Fmoc group:

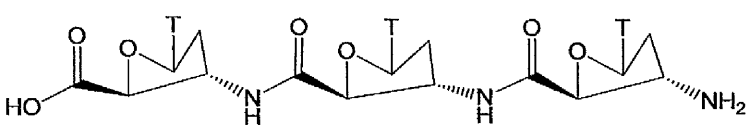
Excess deprotection solution (2% DBU v/v in DMAc) was added to cover the resin and the reaction vessel agitated on an orbital shaker for 10 minutes. The deprotection solution was removed and the resin washed with DMAc (5 x 0.25 mL). A TNBS test was then performed to confirm the removal of the Fmoc protecting group through the detection of the newly generated NH_2 groups.

Further residue couplings:

The cycle of coupling, capping and deprotection was continued according to the above procedures for further nucleoside residues. For all residues apart from the first, coupling times were reduced to 30 minutes and a TNBS test was performed to confirm successful couplings. The quantities for subsequent couplings are shown in Table 1. When all required couplings had been performed and the final deprotection step carried out the resin was dried. Drying of the resin was carried out by washing the resin bed with excess DMAc (x5), DCM (x5) and MeOH (x5) and the application of vacuum for 30 minutes. The resin was then stored in a vacuum dessicator for 24 hours.

Cleavage of the peptide from the solid support:

The dried resin was swollen in excess DCM for 30 minutes and washed with fresh DCM (3 x 0.50 mL). The resin was then covered with the cleavage solution (1% v/v TFA in DCM) and agitated on an orbital shaker for 2 minutes. The cleavage solution was removed and retained. This procedure was repeated a further 9 times. Any residual peptide was washed from the resin bed with excess DCM (x3) and MeOH (x3) then the combined cleavage solutions were evaporated to dryness under reduced pressure to give peptide (**50**).



(50)

When analysed and purified by HPLC, the peptide product (**50**) was found to be contained in those fractions with retention times of 16.39, 17.30 and 18.18 minutes.

HRMS found m/z (ES^-) 728.2299 ($[M-H]^-$ 100%); $[C_{30}H_{34}N_9O_{13}]^-$ requires 728.2276 for peak at retention time 16.39 minutes.

HRMS found m/z (ES^-) 728.2281 ($[M-H]^-$ 100%); $[C_{30}H_{34}N_9O_{13}]^-$ requires 728.2276 for peak at retention time 17.30 minutes.

HRMS found m/z (ES^-) 728.2298 ($[M-H]^-$ 100%); $[C_{30}H_{34}N_9O_{13}]^-$ requires 728.2276 for peak at retention time 18.18 minutes.

Coupling efficiency UV assay:

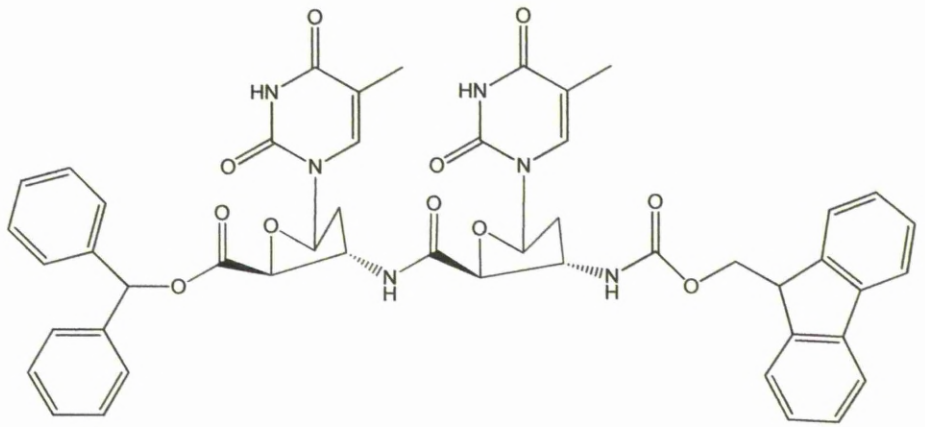
The HMPB-MBHA resin (50 mg) was swollen in DCM overnight and then washed with excess of the chosen solvent DMAc/DMF (x5). Coupling of the first amino acid residue, using selected coupling reagents, was carried out according to the standard procedure discussed previously and all unreacted sites capped. The resin was then washed with excess DMAc/DMF (x5), DCM (x5) and MeOH (x5) and dried overnight in a vacuum dessicator. Once dried the resin was agitated with the deprotection solution (2 mL) for 30 minutes and the solution removed into a 25 mL volumetric flask. This was then made up to the 25 mL mark with MeCN. Aliquots (1 mL) of this solution were then further diluted to 25 mL for assay in the spectrophotometer.

A reference sample was prepared in the same manner with 2 mL of the 2 % DBU solution that had not been in contact with the resin. The UV spectra of these solutions were recorded and the absorption at 294 nm was noted in order to establish the coupling efficiency of a particular combination of resin and reagents.

The concentration of the initial 25 mL solution was then calculated from the UV absorbance and the percentage of coupling could be derived by comparing this to what the expected concentration would be for 100 % coupling.

4.4 Solution synthesis of peptides

Preparation of fully protected thymidine dimer (**51**)



HOBt (118 mg, 0.88 mmol), HBTU (666 mg, 1.76 mmol) and di-*iso*propylethyl amine (230 µl, 1.32 mmol) were dissolved in anhydrous THF (1 mL). 3′-*N*-(9-Fluoromethoxycarbonyl)-3′-deoxythymidyl-5′-carboxylic acid (**3**, 504 mg, 1.05 mmol) dissolved in anhydrous THF (1 mL) was then added to the mixture and the reaction allowed to stir at ambient temperature for 2.5 hours. 3′-Amino-3′-deoxythymidine-5′-benzhydryl ester (**14**, 370 mg, 0.88 mmol) in anhydrous THF (1 mL) was then added to the reactionmixture followed by di-*iso*propylethylamine (230 µl, 1.32 mmol). The reaction was warmed to 40°C and allowed to stir for 1 hour. The reaction was removed from the heat, allowed to cool then all solvents were removed in vacuo to give a brown sticky gum. The crude residue was purified by silica chromatography using a 3-5% methanol/DCM eluent system to give the product as a pale brown foam.

Yield: 688 mg, 0.78 mmol, 89%

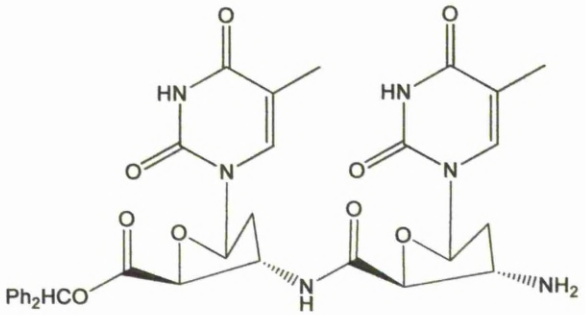
^{1}H NMR (500MHz, Pyr-d_5) δ 1.75 (s, 3H, Thy-CH_3), 1.91 (s, 3H, Thy-CH_3), 2.49 (m, 2H, H2′), 2.60 (m, 2H, H2′), 4.18 (m, 1H, Fmoc-C<u>H</u>), 4.62 (m, 2H, Fmoc-C<u>H</u>$_2$), 4.72 (m, 1H, H4′), 4.92 (m, 1H, H3′), 5.13 (m, 1H, H4′), 5.20 (m, 1H, H3′), 6.68 (m, 1H, H1′), 6.83 (m, 1H, H1′), 7.07-7.77 (m, 18H, Fmoc-H, Benzyhyryl-H), 7.98 (s, 1H, H-5), 8.02 (m, 2H, Fmoc-H), 8.27 (s, 1H, H-5), 9.22 (d, 1H, NH-amide, *J*= 6.50 Hz), 10.03 (d, 1H, NH-amide, *J*= 7.00 Hz), 13.09 (br s, 1H, NH-imine), 13.12 (br s, 1H, NH-imine).

^{13}C NMR (125MHz, Pyr-d_5) δ 12.3 (Thy-CH_3), 12.4 (Thy-CH_3), 36.5 (C2′), 37.2 (C2′), 43.0 (Fmoc-CH), 53.0 (C3′), 55.7 (C3′), 67.0 (Fmoc-CH_2), 78.5 (Benzhydryl-H), 82.2

(C4'), 83.6 (C4'), 86.1 (C1'), 89.5 (C1'), 111.2 (Ar-C), 111.4 (C5), 116.4 (C5), 119.9 (Ar-C), [120.3, 125.0, 126.4, 127.0, 127.2, 127.7, 128.2, 128.3, 128.4 (Ar-C)], 134.7 (C6), 135.9 (C6), 143.6 (C2), 143.7 (C2), 151.0 (C4), 156.4 (C4), 164.3 (Fmoc C=O), 170.0 (C5'), 170.9 (C5').

HRMS found m/z (ES^+) 903.2989 ($[M + Na]^+$ 100%); $[C_{48}H_{44}N_6O_{11} + Na]^+$ requires 903.2966.

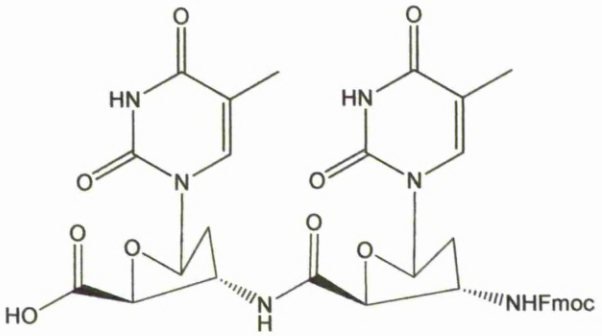
Removal of Fmoc group from fully protected thymidine dimer (**53**)



The fully protected thymidine dimer (**51**, 300 mg, 0.34 mmol) was dissolved in anhydrous THF (3 mL) and DBU (56.0 µl, 0.37 mmol) was added. The reaction was allowed to stir at ambient temperature for 30 mins and all solvents were then removed in vacuo. The crude gum was dissolved in DCM (10 mL) and washed with water (3 x 10 mL). The organic layer was dried over $MgSO_4$ and all solvents removed in vacuo. The crude gum was then redissolved in methanol (10 mL) and washed with hexane (4 x 10 mL) to remove non polar impurities. The impure product was afforded as a pale brown foam.

HRMS found m/z (ES^+) 659.2490 ($[M + H]^+$ 100%); $[C_{33}H_{35}N_6O_9]^+$ requires 659.2466.

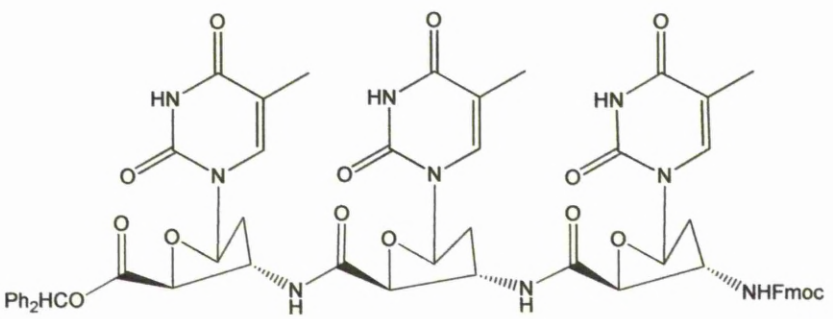
Removal of benzhydryl group from fully protected thymidine dimer (**52**)



The fully protected thymidine dimer (**51**, 150 mg, 0.17 mmol) was dissolved in anhydrous THF (2 mL) and a micro spatula of 10% Pd/C powder was added. The reaction vessel was sealed, evacuated and flushed with nitrogen twice followed hydrogen twice. The reaction was allowed to stir at room temperature and pressure under hydrogen for 72 hours. The system was flushed with fresh nitrogen and then the reaction mixture was filtered through celite. All the solvents were removed in vacuo to give a crude solid which was washed with diethyl ether (10 mL). The solvent was decanted from the solid and this solid redissolved in methanol (10 mL) and washed with hexane (3 x 10 mL). The solvent was removed in vacuo to afford the impure product as a pale brown solid.

HRMS found m/z (ES^+) 737.2171 ($[M + Na]^+$ 100%); $[C_{35}H_{34}N_6O_{11} + Na]^+$ requires 737.2183.

Preparation of fully protected thymidine trimer (**55**)



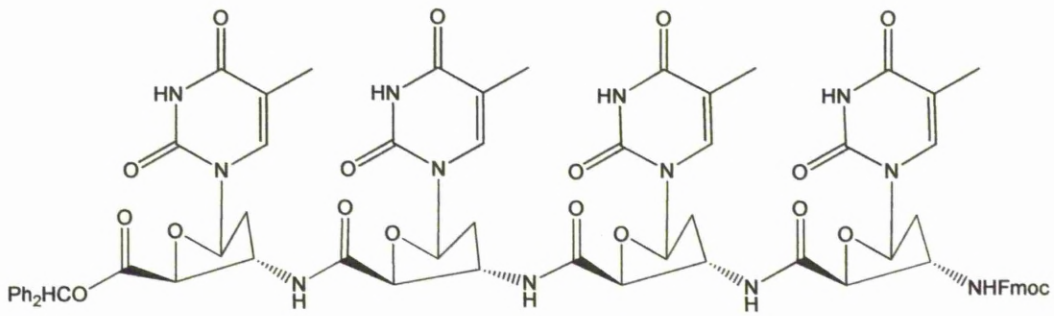
HOBt (8.0 mg, 0.06 mmol), HBTU (45.0 mg, 0.12mmol) and di-*iso*propylethyl amine (15.0 µl, 0.09mmol) were dissolved in anhydrous THF (1mL). 3'-*N*-(9-Fluoromethoxycarbonyl)-3'-deoxythymidyl-5'-carboxylic acid (**3**, 34.0 mg, 0.07 mmol) in anhydrous THF (0.5 mL) was then added and the reaction allowed to stir

at ambient temperature for 2 hours. Dipeptide **53** (39.0 mg, 0.06 mmol) in anhydrous THF (0.5 mL) was then added followed by di-*iso*propylethyl amine (15.0 µl, 0.09 mmol) and the reaction was warmed to 40°C and stirred for 2.5 hours. All solvents were removed in vacuo and the residue was redissolved in a minimum amount of methanol.A precipitate was seen to form as a pale brown solid and was isolated by filteration.

Yield: 11.0 mg, 0.009 mmol, 17%

LRMS found m/z (ES^+) $[M\text{-}C_{13}H_{10} + Na]^+$ 974.1 ($[M + Na]^+$ 100%)

Synthesis of fully protected thymidine tetramer (**54**)

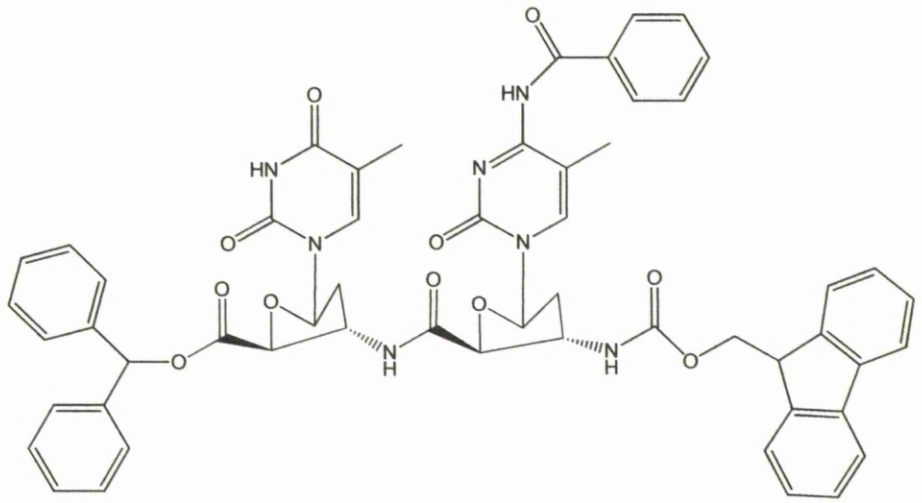


HOBt (10.0 mg, 0.07 mmol), HBTU (55.0 mg, 0.15 mmol) and di-*iso*propylethyl amine (19µl, 0.11 mmol) were dissolved in anhydrous THF (1 mL). Dipeptide **52** (62 mg, 0.09 mmol) in anhydrous THF (0.5 mL) was then added and the reaction allowed to stir at ambient temperature for 2.5 hours. Dipeptide **53** (48 mg, 0.07 mmol) in anhydrous THF (0.5 mL) was then added followed by di-*iso*propylethyl amine (19 µl, 0.11 mmol) and the reaction was warmed to 40°C and allowed to stir for 2.5 hours. All reaction solvents were removed in vacuo and the residue redissolved in a minimum amount of methanol. A precipitate was seen to form as a pale brown solid and was isolated by filteration.

Yield: 20.0 mg, 0.014 mmol, 20%

LRMS found m/z (ES^+) 1377.29 ($[M + Na]^+$ 100%)

Preparation of fully protected thymidine-cytosine dimer (**56**)

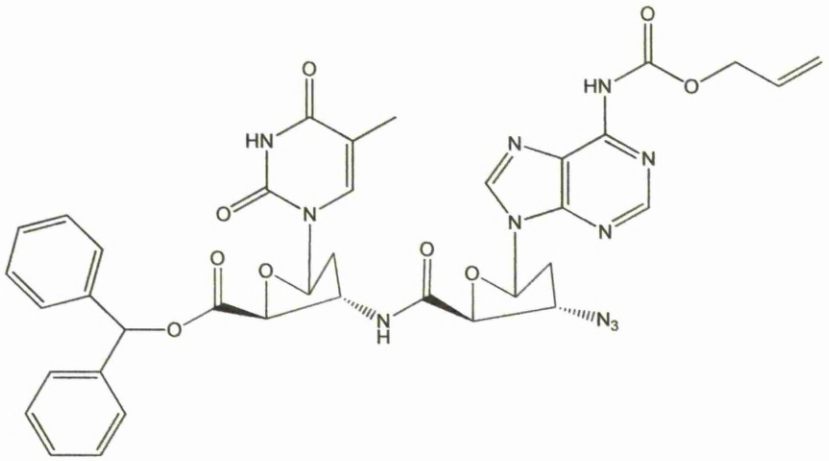


HOBt (8.0 mg, 0.06 mmol), HBTU (44.0 mg, 0.11 mmol) and di-*iso*propylethyl amine (15.0 µl, 0.09 mmol) were dissolved in a mixture of anhydrous THF (1 mL) and anhydrous acetonitrile (0.5 mL). *N*-4-Benzoyl-3′-*N*-Fmoc-amino-2′, 3′-dideoxy-5-methylcytosine-5′-carboxylic acid (**16**, 25.0 mg, 0.07 mmol) in anhydrous THF (0.5 mL) was then added and the reaction mixture allowed to stir at ambient temperature for 2.5 hours. 3′-Amino-3′-deoxythymidine-5′-benzhydryl ester (**14**, 24.0 mg, 0.06 mmol) in anhydrous THF (0.5 mL) was then added followed by di-*iso*propylethyl amine (15 µl, 0.09 mmol) and the reaction was warmed to 40°C and allowed to stir for 2.5 hours. All solvents were removed in vacuo and the crude residue was purified by silica chromatography using a 2-3% methanol/DCM eluent system. The product was afforded as a beige foam.

Yield: 25 mg, 0.025 mmol, 42%

HRMS found m/z (ES^+) 1006.3392 ($[M + Na]^+$ 100%); $[C_{55}H_{49}N_7O_{11} + Na]^+$ requires 1006.3388.

Preparation of fully protected thymidine-adenosine dimer (**57**)



HOBt (21.0 mg, 0.16 mmol), HBTU (118 mg, 0.31 mmol) and di-*iso*propylethyl amine (41.0 µl, 0.23 mmol) were dissolved in a mixture of anhydrous THF (1 mL) and anhydrous acetonitrile (1 mL). *N*-6-Allyloxycarbonyl-3′-azido-2′, 3′-dideoxyadenosine-5′-carboxylic acid (**36**, 70.0 mg, 0.19 mmol) in anhydrous THF (0.5 mL) was then added and the reaction mixture allowed to stir at ambient temperature for 2 hours. 3′-Amino-3′-deoxythymidine-5′-benzhydryl ester (**14**, 65.0 mg, 0.15 mmol) in anhydrous THF (0.5 mL) was then added followed by di-*iso*propylethyl amine (41.0 µl, 0.23 mmol), the reaction was warmed to 40°C and allowed to stir for 4 hours. All reaction solvents were removed in vacuo and the crude residue was analysed by mass spectrometry.

HRMS found m/z (ES^+) 800.2529 ($[M + Na]^+$ 100%); $[C_{37}H_{35}N_{11}O_9 + Na]^+$ requires 800.2517.

Bibliography

1. Portugal, F. H.; Cohen, J. S., *A century of DNA : a history of the discovery of the structure and function of the genetic substance*. M.I.T. Press: Cambridge, Mass. ; London, 1977.
2. Avery, O. T.; McCleod, C. M., *Journal of experimental medicine* **1944,** *79*, 137.
3. Levene, P.; Mori, T., *Journal of Biological Chemistry* **1929,** *81*, 215.
4. Levene, P.; London, E., *Journal of Biological Chemistry* **1929,** *83*, 793.
5. Watson, J.; Crick, F., *Nature* **1953,** *171*, 737.
6. Hoogsteen, K., *Acta Crystallographica* **1963,** *16*, 28.
7. Nikolova, E.; Kim, E.; Wise, A.; O'Brien, P.; Andricioaei, I.; Al-Hashimi, H., *Nature* **2011,** *470*, 498.
8. Altona, C.; Sundaral, M., *Journal of the American Chemical Society* **1972,** *94*, 8205.
9. Olson, W.; Sussman, J., *Journal of the American Chemical Society* **1982,** *104*, 270.
10. Plavec, J.; Tong, W.; Chattopadhyaya, J., *Journal of the American Chemical Society* **1993,** *115*, 9734.
11. Plum, E.; Grollman, A.; Johnson, F.; Breslauer, K., *Biochemistry* **1995,** *34*, 16148.
12. Fabrega, C.; Macias, M.; Eritja, R., *Nucleosides Nucleotides & Nucleic Acids* **2001,** *20*, 251.
13. Hamm, M.; Rajguru, S.; Downs, A.; Cholera, R., *Journal of the American Chemical Society* **2005,** *127*, 12220.
14. Blackburn, G. M., *Nucleic acids in chemistry and biology*. 3rd ed. / edited by G. Michael Blackburn ... [et al.]. ed.; RSC Pub.: Cambridge, 2006.
15. http://foodfashionscience.blogspot.com/2010/08/dna-nanobots-and-origami.html.
16. Dickerson, R., *Scientific American* **1983,** *249*, 94.
17. Dickerson, R.; Drew, H., *Journal of Molecular Biology* **1981,** *149*, 761.
18. Belmont, P.; Constant, J.; Demeunynck, M., *Chemical Society Reviews* **2001,** *30*, 70.
19. Franklin, R.; Gosling, R., *Acta Crystallographica* **1953,** *6*, 673.
20. Fuller, W.; Wilkins, M.; Wilson, H.; Hamilton, L., *Journal of Molecular Biology* **1965,** *12*, 60.
21. Saenger, W., *Principles of nucleic acid structure*. Springer-Verlag: New York, 1984.
22. Lu, X.; Shakked, Z.; Olson, W., *Journal of Molecular Biology* **2000,** *300*, 819.

23. Rich, A.; Zhang, S., *Nature Reviews Genetics* **2003,** *4*, 566.
24. Schroth, G.; Chou, P.; Ho, P., *Journal of Biological Chemistry* **1992,** *267*, 11846.
25. Kim, Y.; Muralinath, M.; Brandt, T.; Pearcy, M.; Hauns, K.; Lowenhaupt, K.; Jacobs, B.; Rich, A., *Proceedings of the National Academy of Sciences of the United States of America* **2003,** *100*, 6974.
26. Urata, H.; Ogura, E.; Shinohara, K.; Ueda, Y.; Akagi, M., *Nucleic Acids Research* **1992,** *20*, 3325.
27. Htun, H.; Dahlberg, J., *Science* **1989,** *243*, 1571.
28. Ghosh, A.; Bansal, M., *Acta Crystallographica Section D-Biological Crystallography* **2003,** *59*, 620.
29. Xu, Y., *Chemical Society Reviews* **2011,** *40*, 2719.
30. Brown, R.; Hurley, L., *Biochemical Society Transactions* **2011,** *39*, 635.
31. Burge, S.; Parkinson, G.; Hazel, P.; Todd, A.; Neidle, S., *Nucleic Acids Research* **2006,** *34*, 5402.
32. Ladame, S.; Schouten, J.; Stuart, J.; Roldan, J.; Neidle, S.; Balasubramanian, S., *Organic & Biomolecular Chemistry* **2004,** *2*, 2925.
33. Mergny, J.; Riou, J.; Mailliet, P.; Teulade-Fichou, M.; Gilson, E., *Nucleic Acids Research* **2002,** *30*, 839.
34. Fernando, H.; Sewitz, S.; Darot, J.; Tavare, S.; Huppert, J.; Balasubramanian, S., *Nucleic Acids Research* **2009,** *37*, 6716.
35. Gehring, K.; Leroy, J.; Gueron, M., *Nature* **1993,** *363*, 561.
36. Mergny, J.; Lacroix, L.; Han, X.; Leroy, J.; Helene, C., *Journal of the American Chemical Society* **1995,** *117*, 8887.
37. Catasti, P.; Chen, X.; Deaven, L.; Moyzis, R.; Bradbury, E.; Gupta, G., *Journal of Molecular Biology* **1997,** *272*, 369.
38. Han, X.; Leroy, J.; Gueron, M., *Journal of Molecular Biology* **1998,** *278* (5), 949.
39. Kendrick, S.; Akiyama, Y.; Hecht, S.; Hurley, L., *Journal of the American Chemical Society* **2009,** *131*, 17667.
40. Alberti, P.; Bourdoncle, A.; Sacca, B.; Lacroix, L.; Mergny, J., *Organic & Biomolecular Chemistry* **2006,** *4*, 3383.
41. Campolongo, M.; Kahn, J.; Cheng, W.; Yang, D.; Gupton-Campolongo, T.; Luo, D., *Journal of Materials Chemistry* **2011,** *21*, 6113.
42. Kaucher, M.; Harrell, W.; Davis, J., *Journal of the American Chemical Society* **2006,** *128*, 38.

43. Li, T.; Wang, E.; Dong, S., *Journal of the American Chemical Society* **2009,** *131*, 15082.
44. Cavalieri, F.; Ng, S.; Mazzuca, C.; Jia, Z.; Bulmus, V.; Davis, T.; Caruso, F., *Small* **2011,** *7*, 101.
45. Song, G.; Chen, M.; Chen, C.; Wang, C.; Hu, D.; Ren, J.; Qu, X., *Biochimie* **2010,** *92*, 121.
46. Neilsen, P.; Egholm, M.; Berg, R.; Buchardt, O., *Science* **1991,** *254*, 1497.
47. Hyrup, B.; Nielsen, P., *Bioorganic & Medicinal Chemistry* **1996,** *4*, 5.
48. Ray, A.; Norden, B., *Faseb Journal* **2000,** *14*, 1041.
49. Good, L.; Nielsen, P., *Nature Biotechnology* **1998,** *16*, 355.
50. Ma, Z.; Olechnowicz, F.; Skorik, Y.; Achim, C., *Inorganic Chemistry* **2011,** *50*, 6083.
51. Nielsen, P., *Chemistry & Biodiversity* **2010,** *7*, 786.
52. Dueholm, K.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H.; Vulpius, T.; Petersen, K.; Berg, R.; Nielsen, P.; Buchardt, O., *Journal of Organic Chemistry* **1994,** *59*, 5767.
53. Christensen, L.; Fitzpatrick, R.; Glidea, B. D.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P.; Coull, J. M.; Berg, J. M., *Journal of Peptide Science* **1995,** *1*, 175.
54. Porcheddu, A.; Giacomelli, G., *Current Medicinal Chemistry* **2005,** *12*, 2561.
55. Kumar, V.; Ganesh, K., *Accounts of Chemical Research* **2005,** *38*, 404.
56. Demidov, V.; Potaman, V.; Frankkamenetskii, M.; Egholm, M.; Buchardt, O.; Sonnichsen, S.; Nielsen, P., *Biochemical Pharmacology* **1994,** *48*, 1310.
57. Demidov, V.; Frankkamenetskii, M.; Egholm, M.; Buchardt, O.; Nielsen, P., *Nucleic Acids Research* **1993,** *21*, 2103.
58. Hyrup, B.; Egholm, M.; Nielsen, P.; Wittung, P.; Norden, B.; Buchardt, O., *Journal of the American Chemical Society* **1994,** *116*, 7964.
59. Nielsen, P.; Egholm, M.; Buchardt, O., *Bioconjugate Chemistry* **1994,** *5*, 3.
60. Zhou, P.; Wang, M.; Du, L.; Fisher, G.; Waggoner, A.; Ly, D., *Journal of the American Chemical Society* **2003,** *125*, 6878.
61. Gangamani, B.; Kumar, V.; Ganesh, K., *Biochemical and Biophysical Research Communications* **1997,** *240*, 778.
62. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D., *Angewandte Chemie-International Edition* **1998,** *37*, 2797.
63. Eriksson, M.; Nielsen, P., *Nature Structural Biology* **1996,** *3*, 410.
64. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S.; Driver, D.; Berg, R.; Kim, S.; Norden, B.; Neilsen, P., *Nature* **1993,** *365*, 566.

65. Jensen, K.; Orum, H.; Nielsen, P.; Norden, B., *Biochemistry* **1997,** *36*, 5072.
66. Rasmussen, H.; Sandholm, J., *Nature Structural Biology* **1997,** *4*, 98.
67. Nielsen, P.; Haaima, G., *Chemical Society Reviews* **1997,** *26*, 73.
68. Nielsen, P., *Accounts of Chemical Research* **1999,** *32* (7), 624.
69. Bonham, M.; Brown, S.; Boyd, A.; Brown, P.; Bruckenstein, D.; Hanvey, J.; Thomson, S.; Pipe, A.; Hassman, F.; Bisi, J.; Froehler, B.; Matteucci, M.; Wagner, R.; Noble, S.; Babiss, L., *Nucleic Acids Research* **1995,** *23*, 1197.
70. Pokorski, J.; Witschi, M.; Purnell, B.; Appella, D., *Journal of the American Chemical Society* **2004,** *126*, 15067.
71. Vilaivan, T.; Srisuwannaket, C., *Organic Letters* **2006,** *8*, 1897.
72. Lagriffoule, P.; Wittung, P.; Eriksson, M.; Jensen, K.; Norden, B.; Buchardt, O.; Nielsen, P., *Chemistry-a European Journal* **1997,** *3*, 912.
73. Pauling, L.; Corey, R.; Branson, H., *Proceedings of the National Academy of Sciences of the United States of America* **1951,** *37*, 205.
74. Astbury, W. T.; Street, A., *Transactions of the Royal Society (London)* **1931,** *A230*, 75.
75. Neurath, H., *Journal of Physical Chemistry* **1940,** *44*, 296.
76. Pauling, L.; Corey, R., *Nature* **1951,** *168*, 550.
77. Eisenberg, D., *Proceedings of the National Academy of Sciences of the United States of America* **2003,** *100*, 11207.
78. Dunitz, J., *Angewandte Chemie-International Edition* **2001,** *40*, 4167.
79. Fodje, M.; Al-Karadaghi, S., *Protein Engineering* **2002,** *15*, 353.
80. Huggins, M., *Chemical Reviews* **1943,** *32*, 195.
81. Hatakeyama, Y.; Sawada, T.; Kawano, M.; Fujita, M., *Angewandte Chemie-International Edition* **2009,** *48*, 8695.
82. Riek, R.; Graham, R., *Journal of Structural Biology* **2011,** *173* (1), 153.
83. Chou, P.; Fasman, G., *Biochemistry* **1974,** *13*, 222.
84. Pace, C.; Scholtz, J., *Biophysical Journal* **1998,** *75*, 422.
85. Frishman, D.; Argos, P., *Folding & Design* **1997,** *2*, 159.
86. Cortajarena, A.; Mochrie, S.; Regan, L., *Protein Science* **2011,** *20*, 1042.
87. Pauling, L.; Corey, R., *Proceedings of the National Academy of Sciences of the United States of America* **1951,** *37*, 251.
88. Sarroukh, R.; Cerf, E.; Derclaye, S.; Dufrene, Y.; Goormaghtigh, E.; Ruysschaert, J.; Raussens, V., *Cellular and Molecular Life Sciences* **2011,** *68*, 1429.
89. Branden, C.; Tooze, J., *Introduction to protein structure*. 2nd ed.; Garland: New York, 1999.

90. Voet, D.; Voet, J. G., *Biochemistry*. 3rd ed.; Wiley: Hoboken, N J, 2004.
91. Garrett, R.; Grisham, C. M., *Biochemistry*. 4th , International ed.; Brooks/Cole ; Cengage Learning: Australia ; United Kingdom, 2010.
92. Nesloney, C.; Kelly, J., *Bioorganic & Medicinal Chemistry* **1996,** *4*, 739.
93. Nowick, J.; Smith, E.; Ziller, J.; Shaka, A., *Tetrahedron* **2002,** *58*, 727.
94. Gellman, S., *Accounts of Chemical Research* **1998,** *31*, 173.
95. Porter, E.; Wang, X.; Lee, H.; Weisblum, B.; Gellman, S., *Nature* **2000,** *404*, 565.
96. Perez-Alvite, M. J.; Mosquera, M.; Castedo, L.; Granja, J. R., *Amino acids* **2011,** *41*, 621.
97. Gutiérrez-Abad, R.; Carbajo, D.; Nolis, P.; Acosta-Silva, C.; Cobos, J.; Illa, O.; Royo, M.; Ortuño, R., *Amino Acids* **2011,** *41*, 673.
98. Jiang, H.; Leger, J.; Huc, I., *Journal of the American Chemical Society* **2003,** *125*, 3448.
99. Simone, M.; Edwards, A.; Tranter, G.; Fleet, G., *Amino Acids* **2011,** *41*, 643.
100. Cornelissen, J. J. L. M.; Graswinckel, W. S.; Rowan, A. E.; Sommerdijk, N. A. J. M.; Nolte, R. J. M., *Journal of Polymer Science Part A: Polymer Chemistry* **2003,** *41*, 1725.
101. Gojkovic, Z.; Sandrini, M.; Piskur, J., *Genetics* **2001,** *158*, 999.
102. Cheng, R.; Gellman, S.; DeGrado, W., *Chemical Reviews* **2001,** *101*, 3219.
103. Avignon, M.; Huong, P.; Lascombe, J.; Marraud, M.; Neel, J., *Biopolymers* **1969,** *8*, 69.
104. Dado, G.; Gellman, S., *Journal of the American Chemical Society* **1994,** *116*, 1054.
105. Bestian, H., *Angewandte Chemie-International Edition* **1968,** *7*, 278.
106. Lengyel, G.; Frank, R.; Horne, W., *Journal of the American Chemical Society* **2011,** *133*, 4246.
107. Barchi, J.; Huang, X.; Appella, D.; Christianson, L.; Durell, S.; Gellman, S., *Journal of the American Chemical Society* **2000,** *122*, 2711.
108. Muller, H.; Seebach, D., *Angewandte Chemie-International Edition in English* **1993,** *32*, 477.
109. Seebach, D.; Overhand, M.; Kuhnle, F.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H., *Helvetica Chimica Acta* **1996,** *79*, 913.
110. Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D., *Helvetica Chimica Acta* **1996,** *79*, 2043.
111. Seebach, D.; Abele, S.; Gademann, K.; Guichard, G.; Hintermann, T.; Jaun, B.; Matthews, J.; Schreiber, J., *Helvetica Chimica Acta* **1998,** *81*, 932.

112. O'Neil, K.; Degrado, W., *Science* **1990,** *250*, 646.
113. Glickson, J.; Applequi, J., *Journal of the American Chemical Society* **1971,** *93*, 3276.
114. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H., *Journal of the American Chemical Society* **1998,** *118*, 13071.
115. Appella, D.; Christianson, L.; Klein, D.; Richards, M.; Powell, D.; Gellman, S., *Journal of the American Chemical Society* **1999,** *121*, 7574.
116. Claridge, T.; Goodman, J.; Moreno, A.; Angus, D.; Barker, S.; Taillefumier, C.; Watterson, M.; Fleet, G., *Tetrahedron Letters* **2001,** *42*, 4251.
117. Fernandes, C.; Faure, S.; Pereira, E.; Thery, V.; Declerck, V.; Guillot, R.; Aitken, D., *Organic Letters* **2010,** *12*, 3606.
118. Abele, S.; Seiler, P.; Seebach, D., *Helvetica Chimica Acta* **1999,** *82*, 1559.
119. Torres, E.; Gorrea, E.; Burusco, K.; Da Silva, E.; Nolis, P.; Rua, F.; Boussert, S.; Diez-Perez, I.; Dannenberg, S.; Izquierdo, S.; Giralt, E.; Jaime, C.; Branchadell, V.; Ortuno, R., *Organic & Biomolecular Chemistry* **2010,** *8*, 564.
120. Torres, E.; Gorrea, E.; Da Silva, E.; Nolis, P.; Branchadell, V.; Ortuno, R., *Organic Letters* **2009,** *11*, 2301.
121. Seebach, D.; Abele, S.; Sifferlen, T.; Hänggi, M.; Gruner, S.; Seiler, P., *Helvetica Chimica Acta* **1998,** *81*, 2218.
122. Seebach, D.; Gademann, K.; Schreiber, J.; Matthews, J.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H., *Helvetica Chimica Acta* **1997,** *80*, 2033.
123. Wu, Y.; Wang, D., *Journal of the American Chemical Society* **1999,** *121*, 9352.
124. Appella, D.; Barchi, J.; Durell, S.; Gellman, S., *Journal of the American Chemical Society* **1999,** *121*, 2309.
125. Porter, E.; Wang, X.; Schmitt, M.; Gellman, S., *Organic Letters* **2002,** *4*, 3317.
126. Krauthauser, S.; Christianson, L.; Powell, D.; Gellman, S., *Journal of the American Chemical Society* **1997,** *119*, 11719.
127. Chung, Y.; Christianson, L.; Stanger, H.; Powell, D.; Gellman, S., *Journal of the American Chemical Society* **1998,** *120*, 10555.
128. Seebach, D.; Abele, S.; Gademann, K.; Jaun, B., *Angewandte Chemie-International Edition* **1999,** *38*, 1595.
129. Lelais, G.; Seebach, D.; Jaun, B.; Mathad, R.; Flogel, O.; Rossi, F.; Campo, M.; Wortmann, A., *Helvetica Chimica Acta* **2006,** *89*, 361.

130. Seebach, D.; Matthews, J.; Meden, A.; Wessels, T.; Baerlocher, C.; McCusker, L., *Helvetica Chimica Acta* **1997,** *80*, 173.
131. Ishihara, Y.; Kimura, S., *Journal of Peptide Science* **2010,** *16*, 110.
132. Brazeau, P.; Vale, W.; Burgus, R.; Ling, N.; Butcher, M.; Rivier, J.; Guillemi, R., *Science* **1973,** *179*, 77.
133. Melacini, G.; Zhu, Q.; Osapay, G.; Goodman, M., *Journal of Medicinal Chemistry* **1997,** *40*, 2252.
134. Gademann, K.; Ernst, M.; Hoyer, D.; Seebach, D., *Angewandte Chemie-International Edition* **1999,** *38*, 1223.
135. Horne, W.; Price, J.; Keck, J.; Gellman, S., *Journal of the American Chemical Society* **2007,** *129*, 4178.
136. Price, J.; Horne, W.; Gellman, S., *Journal of the American Chemical Society* **2007,** *129*, 6376.
137. Craig, C.; Goodman, J.; Schepartz, A., *Chembiochem* **2011,** *12*, 1035.
138. Daniels, D.; Petersson, E.; Qiu, J.; Schepartz, A., *Journal of the American Chemical Society* **2007,** *129*, 1532.
139. Chakraborty, T.; Kumar, S.; Mohan, B.; Sarma, G.; Kiran, M.; Jagadeesh, B., *Tetrahedron Letters* **2007,** *48*, 6945.
140. Chakraborty, T.; Ghosh, S.; Jayaprakash, S., *Current Medicinal Chemistry* **2002,** *9*, 421.
141. Gruner, S.; Keri, G.; Schwab, R.; Venetianer, A.; Kessler, H., *Organic Letters* **2001,** *3*, 3723.
142. Claridge, T.; Long, D.; Baker, C.; Odell, B.; Grant, G.; Edwards, A.; Tranter, G.; Fleet, G.; Smith, M., *Journal of Organic Chemistry* **2005,** *70*, 2082.
143. Szabo, L.; Smith, B.; McReynolds, K.; Parrill, A.; Morris, E.; Gervay, J., *Journal of Organic Chemistry* **1998,** *63*, 1074.
144. McReynolds, K.; Gervay-Hague, J., *Tetrahedron-Asymmetry* **2000,** *11*, 337.
145. Risseeuw, M.; Overhand, M.; Fleet, G.; Simone, M., *Tetrahedron-Asymmetry* **2007,** *18*, 2001.
146. Andreini, M.; Taillefumier, C.; Chretien, F.; Thery, V.; Chapleur, Y., *Journal of Organic Chemistry* **2009,** *74*, 7651.
147. Smith, M.; Claridge, T.; Sansom, M.; Fleet, G., *Organic & Biomolecular Chemistry* **2003,** *1*, 3647.
148. Doel, M. T.; Jones, A. S.; Taylor, N., *Tetrahedron Letters* **1969,** *10*, 2285.
149. Yamazaki, T.; Komatsu, K.; Umemiya, H.; Hashimoto, Y.; Shudo, K.; Kagechika, H., *Tetrahedron Letters* **1997,** *38*, 8363.

150. Brückner, Arndt M.; Garcia, M.; Marsh, A.; Gellman, Samuel H.; Diederichsen, U., *European Journal of Organic Chemistry* **2003,** *2003*, 3555.
151. Weiß, A.; Diederichsen, U., *European Journal of Organic Chemistry* **2007,** *2007*, 5531.
152. Brückner, A. M.; Chakraborty, P.; Gellman, S. H.; Diederichsen, U., *Angewandte Chemie International Edition* **2003,** *42*, 4395.
153. Threlfall, R.; Davies, A.; Howarth, N. M.; Fisher, J.; Cosstick, R., *Chemical. Communications* **2008**, 585.
154. Chandrasekhar, S.; Kiranmai, N.; Kiran, M. U.; Devi, A. S.; Reddy, G. P. K.; Idris, M.; Jagadeesh, B., *Chemical. Communications* **2010,** *46*, 6962.
155. Threlfall, R., PhD thesis: Design and synthesis of foldamers derived from nucleoside beta-amino acids. 2008.
156. Threlfall, R.; Davies, A.; Howarth, N.; Cosstick, R., *Nucleosides, Nucleotides and Nucleic Acids* **2007,** *26*, 611.
157. Chandrasekhar, S.; Reddy, G.; Kiran, M.; Nagesh, C.; Jagadeesh, B., *Tetrahedron Letters* **2008,** *49*, 2969.
158. Varma, R. S.; Hogan, M. E., *Tetrahedron Letters* **1992,** *33*, 7719.
159. Pathak, T., *Chemical Reviews* **2002,** *102*, 1623.
160. Fleet, G.; Son, J.; Derome, A., *Tetrahedron* **1988,** *44*, 625.
161. Jung, M.; Gardiner, J., *Journal of Organic Chemistry* **1991,** *56*, 2614.
162. Hager, M.; Liotta, D., *Journal of the American Chemical Society* **1991,** *113*, 5117.
163. Agyeiaye, K.; Yan, S.; Hebbler, A.; Baker, D., *Nucleosides & Nucleotides* **1989,** *8*, 327.
164. Czernecki, S.; Valéry, J.-M., *Synthesis* **1991,** *1991*, 239.
165. Maillard, M.; Faraj, A.; Frappier, F.; Florent, J.; Grierson, D.; Monneret, C., *Tetrahedron Letters* **1989,** *30*, 1955.
166. Horwitz, J.; Chua, J.; Noel, M., *Journal of Organic Chemistry* **1964,** *29*, 2076.
167. Lin, T.; Prusoff, W., *Journal of Medicinal Chemistry* **1978,** *21*, 109.
168. Scriven, E.; Turnbull, K., *Chemical Reviews* **1988,** *88*, 297.
169. Nooy, A. E. J. d.; Besemer, A. C.; Bekkum, H. v., *Synthesis* **1996,** *1996*, 1153.
170. Vogler, T.; Studer, A., *Synthesis* **2008,** *2008*, 1979.
171. Bobbitt, J.; Flores, M., *Heterocycles* **1988,** *27*, 509.
172. Cella, J.; Kelley, J.; Kenehan, E., *Journal of Organic Chemistry* **1975,** *40*, 1860.

173. Miyazawa, T.; Endo, T.; Shiihashi, S.; Okawara, M., *Journal of Organic Chemistry* **1985,** *50*, 1332.
174. Miyamoto, K.; Tada, N.; Ochiai, M., *Journal of the American Chemical Society* **2007,** *129*, 2772.
175. Flyunt, R.; Theruvathu, J.; Leitzke, A.; von Sonntag, C., *Journal of the Chemical Society-Perkin Transactions 2* **2002**, 1572.
176. Lapatsanis, L., *Tetrahedron Letters* **1978**, 4697.
177. Stelakat, G.; Paganou, A.; Zervas, L., *Journal of the Chemical Society C-Organic* **1966**, 1191.
178. Javed, M.; Brewer, M., *Organic Letters* **2007,** *9*, 1789.
179. Javed, M. I.; Brewer, M., *Organic Synthesis* **2008,** *85*, 189.
180. Curini, M.; Rosati, O.; Pisani, E.; Cabri, W.; Brusco, S.; Riscazzi, M., *Tetrahedron Letters* **1997,** *38*, 1239.
181. Greenhalgh, R., *Synlett* **1992**, 235.
182. Boyer, J., *Journal of the American Chemical Society* **1951,** *73*, 5865.
183. Benati, L.; Bencivenni, G.; Leardini, R.; Nanni, D.; Minozzi, M.; Spagnolo, P.; Scialpi, R.; Zanardi, G., *Organic Letters* **2006,** *8*, 2499.
184. Zheng, P.; Duan, X.; Wang, W.; Wang, X.; Guo, Y.; Tu, Q., *Chinese Journal of Chemistry* **2006,** *24*, 825.
185. Li, J. J.; Corey, E. J., *Name reactions for functional group transformations*. Wiley ; Chichester : John Wiley [distributor]: Hoboken, N.J., 2007.
186. Robins, M. J.; Doboszewski, B.; Nilsson, B. L.; Peterson, M. A., *Nucleosides, Nucleotides and Nucleic Acids* **2000,** *19*, 69.
187. Gogoi, K.; Gunjal, A. D.; Phalgune, U. D.; Kumar, V. A., *Organic Letters* **2007,** *9*, 2697.
188. Fox, J. J.; Van Praag, D.; Wempen, I.; Doerr, I. L.; Cheong, L.; Knoll, J. E.; Eidinoff, M. L.; Bendich, A.; Brown, G. B., *Journal of the American Chemical Society* **1959,** *81*, 178.
189. Münzel, M.; Globisch, D.; Brückl, T.; Wagner, M.; Welzmiller, V.; Michalakis, S.; Müller, M.; Biel, M.; Carell, T., *Angewandte Chemie International Edition* **2010,** *49*, 5375.
190. Busson, R.; Kerremans, L.; Van Aerschot, A.; Peeters, M.; Blaton, N.; Herdewijn, P., *Nucleosides and Nucleotides* **1999,** *18*, 1079.
191. González-Moa, M. J.; Besada, P.; Terán, C.; Santana, L.; Uriarte, E.; Viña, D., *Helvetica Chimica Acta* **2006,** *89*, 954.
192. Sung, W. L., *Journal of the Chemical Society, Chemical Communications* **1981**, 1089.

193. Bergmann, F.; Bannwarth, W.; Tam, S., *Tetrahedron Letters* **1995,** *36*, 6823.
194. Gaynor, J.; Cosstick, R., *Current Organic Chemistry* **2008,** *12*, 291.
195. Hutchison, A.; Williams, M.; Dejesus, R.; Yokoyama, R.; Oei, H.; Ghai, G.; Webb, R.; Zoganas, H.; Stone, G.; Jarvis, M., *Journal of Medicinal Chemistry* **1990,** *33*, 1919.
196. Barton, D.; Gero, S.; Quicletsire, B.; Samadi, M., *Tetrahedron Letters* **1989,** *30*, 4969.
197. Zhang, J.; Matteucci, M., *Bioorganic & Medicinal Chemistry Letters* **1999,** *9*, 2213.
198. Rejzek, M.; Mukhopadhyay, B.; Wenzel, C.; Lam, J.; Field, R., *Carbohydrate Research* **2007,** *342*, 460.
199. Mancuso, A.; Huang, S.; Swern, D., *Journal of Organic Chemistry* **1978,** *43*, 2480.
200. Thompson, S.; Heathcock, C., *Journal of Organic Chemistry* **1992,** *57*, 5979.
201. Albright, J.; Goldman, L., *Journal of the American Chemical Society* **1967,** *89*, 2416.
202. Omura, K.; Swern, D., *Tetrahedron* **1978,** *34*, 1651.
203. Pfitzner, K.; Moffatt, J., *Journal of the American Chemical Society* **1965,** *87*, 5661.
204. Pfitzner, K.; Moffatt, J., *Journal of the American Chemical Society* **1965,** *87*, 5670.
205. Dess, D.; Martin, J., *Journal of the American Chemical Society* **1991,** *113*, 7277.
206. Meyer, S.; Schreiber, S., *Journal of Organic Chemistry* **1994,** *59*, 7549.
207. Herdewijn, P., *Journal of Organic Chemistry* **1988,** *53*, 5050.
208. Heidenhain, S.; Hayakawa, Y., *Synlett* **1998**, 853.
209. Markiewicz, W. T.; Wiewiórowski, M., *Nucleic Acids Research* **1978,** *1*, s185.
210. Markiewicz, W., *Journal of Chemical Research* **1979**, 24.
211. Sigmund, H.; Pfleiderer, W., *Helvetica Chimica Acta* **2003,** *86*, 2299.
212. Hayakawa, Y.; Wakabayashi, S.; Kato, H.; Noyori, R., *Journal of the American Chemical Society* **1990,** *112*, 1691.
213. Samano, V.; Robins, M. J., *The Journal of Organic Chemistry* **1990,** *55*, 5186.
214. Hansske, F.; Madej, D.; Robins, M. J., *Tetrahedron* **1984,** *40*, 125.
215. Eisenhuth, R.; Richert, C., *The Journal of Organic Chemistry* **2008,** *74*, 26.
216. Ikehara, M.; Nakahara, Y.; Yamada, S., *Chem. Pharm. Bull.* **1971,** *19*, 538.

217. Robins, M. J.; Nowak, I.; Wnuk, S. F.; Hansske, F.; Madej, D., *Journal of Organic Chemistry* **2007,** *72*, 8216.
218. Lavandera, I.; Fernández, S.; Ferrero, M.; Gotor, V., *Tetrahedron* **2003,** *59*, 5449.
219. Merrifield, R. B., *Journal of the American Chemical Society* **1963,** *85*, 2149.
220. Merrifield, R. B., *Journal of the American Chemical Society* **1964,** *86*, 304.
221. Carpino, L. A.; Han, G. Y., *The Journal of Organic Chemistry* **1972,** *37*, 3404.
222. Wang, S.-S., *Journal of the American Chemical Society* **1973,** *95*, 1328.
223. Rink, H., *Tetrahedron Letters* **1987,** *28*, 3787.
224. Buskas, T.; Ingale, S.; Boons, G.-J., *Angewandte Chemie International Edition* **2005,** *44*, 5985.
225. Sieber, P., *Tetrahedron Letters* **1987,** *28*, 2107.
226. Han, S.-Y.; Kim, Y.-A., *Tetrahedron* **2004,** *60*, 2447.
227. König, W.; Geiger, R., *Chemische Berichte* **1970,** *103*, 788.
228. Carpino, L. A.; El-Faham, A., *Tetrahedron* **1999,** *55*, 6813.
229. Boger, D. L.; Miyazaki, S.; Kim, S. H.; Wu, J. H.; Castle, S. L.; Loiseleur, O.; Jin, Q., *Journal of the American Chemical Society* **1999,** *121*, 10004.
230. Izdebski, J.; Pachulska, M.; Orlowska, A., *International Journal of Peptide and Protein Research.* **1994,** *44*, 414.
231. Castro, B.; Dormoy, J. R., *Tetrahedron Letters* **1972**, 4747.
232. Castro, B.; Dormoy, J. R., *Tetrahedron Letters* **1973**, 3243.
233. Castro, B.; Dormoy, J.-R.; Dourtoglou, B.; Evin, G.; Selve, C.; Ziegler, J.-C., *Synthesis* **1976**, 751.
234. Taunton, J.; Collins, J. L.; Schreiber, S. L., *Journal of the American Chemical Society* **1996,** *118*, 10412.
235. Coste, J.; Le-Nguyen, D.; Castro, B., *Tetrahedron Letters* **1990,** *31*, 205.
236. Tercel, M.; Atwell, G. J.; Yang, S.; Ashoorzadeh, A.; Stevenson, R. J.; Botting, K. J.; Gu, Y.; Mehta, S. Y.; Denny, W. A.; Wilson, W. R.; Pruijn, F. B., *Angew. Chem., Int. Ed.* **2011,** *50*, 2606.
237. Carpino, L. A. New reagents for peptide couplings. WO9407910A1, 1994.
238. Palecek, J.; Draeger, G.; Kirschning, A., *Synthesis* **2011**, 653.
239. Dourtoglou, V.; Ziegler, J. C.; Gross, B., *Tetrahedron Letters.* **1978**, 1269.
240. Avadisian, M.; Fletcher, S.; Liu, B.; Zhao, W.; Yue, P.; Badali, D.; Xu, W.; Schimmer, A. D.; Turkson, J.; Gradinaru, C. C.; Gunning, P. T., *Angew. Chem., Int. Ed.* **2011,** *50*, 6248.

241. Carpino, L. A.; Imazumi, H.; El-Faham, A.; Ferrer, F. J.; Zhang, C.; Lee, Y.; Foxman, B. M.; Henklein, P.; Hanay, C.; Mugge, C.; Wenschuh, H.; Klose, J.; Beyermann, M.; Bienert, M., *Angewandte Chemie International Edition* **2002,** *41*, 441.

242. Aaronson, J. G.; Klein, L. J.; Momose, A. A.; O'Brien, A. M.; Shaw, A. W.; Tucker, T. J.; Yuan, Y.; Tellers, D. M., *Bioconjugate Chem.* **2011,** *22*, 1723.

243. Albericio, F.; Bofill, J. M.; El-Faham, A.; Kates, S. A., *Journal of Organic Chemistry* **1998,** *63*, 9678.